

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-537760

(P2004-537760A)

(43) 公表日 平成16年12月16日(2004.12.16)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
G 1 O L 11/00	G 1 O L 9/00 A	5 D O 1 5
G 1 O L 15/00	G 1 O L 9/00 E	
G 1 O L 15/10	G 1 O L 3/00 5 3 1 N	
	G 1 O L 3/00 5 5 1 A	

審査請求 未請求 予備審査請求 未請求 (全 56 頁)

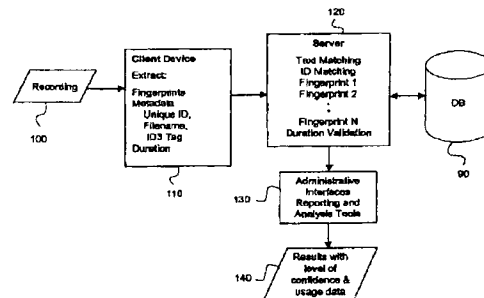
(21) 出願番号	特願2003-517800 (P2003-517800)	(71) 出願人	504040988
(86) (22) 出願日	平成14年7月31日 (2002.7.31)		グレースノート インコーポレイテッド
(85) 翻訳文提出日	平成16年1月30日 (2004.1.30)		アメリカ合衆国 カリフォルニア州 9 4
(86) 国際出願番号	PCT/US2002/024054		7 1 0 バークレイ 4 ストリート 2 1
(87) 国際公開番号	W02003/012695		2 4
(87) 国際公開日	平成15年2月13日 (2003.2.13)	(74) 代理人	100070150
(31) 優先権主張番号	60/308,594		弁理士 伊東 忠彦
(32) 優先日	平成13年7月31日 (2001.7.31)	(74) 代理人	100094525
(33) 優先権主張国	米国 (US)		弁理士 土井 健二
		(74) 代理人	100094514
			弁理士 林 恒徳
		(72) 発明者	デイル ティー ロバーツ
			アメリカ合衆国 カリフォルニア州 9 4
			9 6 0 サンアンセルモ オークスプリング
			ドライブ 1 5

最終頁に続く

(54) 【発明の名称】 録音の多段階識別関連出願の相互参照本出願は、「デジタル音楽の多段階識別の方法およびシステム」と題された米国仮出願第60/308,594号 (発明者: デールT. ロバーツ (D a l

(57) 【要約】

未識別録音から複数の情報およびそれに関連付けられた情報が抽出される。関連付けられた情報としては、録音が例えばMP3フォーマット等のコンピュータ・ファイルである場合はファイル名が、また録音がコンパクトディスク等のリムーバブルメディアにある場合は目次(TOC)データが含まれる。少なくとも1つ、好適には複数の音紋がアルゴリズムにより決定され、一種以上の音紋抽出方法を用いて録音から抽出される。抽出された情報は、基準録音用に維持されているデータベースの対応する情報と比較される。識別は、例えばハッシュID、一意ID、またはテキスト情報等の利用可能な最も正確かつ効率的な方法を用いて開始される。音紋マッチングが他のマッチング結果を裏付けるのに用いられ、未識別録音と合致見込み基準録音の演奏時間同士を比較することにより検証が実行される。



【特許請求の範囲】

【請求項 1】

録音を識別する方法であって、
ユーザー所有の媒体に保存された未識別録音に関する情報と、前記未識別録音の少なくとも 1 つの部分から少なくとも 1 つのアルゴリズムにより決定された音紋とを抽出するステップと、
前記未識別録音から抽出された情報の少なくとも一部と、基準録音の対応する情報を含む識別データベースとを用いて、前記未識別録音の見込み識別を決定するステップと、
前記少なくとも 1 つのアルゴリズムにより決定された音紋と組み合わせられた前記情報の少なくとも一部の各々に基づく前記見込み識別が、単一の基準録音を各々の信用度で識別した場合に前記未識別録音を識別するステップと、を含む方法。

10

【請求項 2】

前記識別データベースが識別サービス・プロバイダにより管理されており、
前記決定ステップが、前記未識別録音に前記識別サービス・プロバイダから一意な識別子が関連付けられている場合は前記一意識別子を用い、又は前記未識別録音にテキストが関連付けられている場合は前記未識別録音に関連付けられている前記テキストを用いる、請求項 1 に記載の方法。

【請求項 3】

前記未識別録音の抽出された長さを前記単一基準録音の保存された長さと比較することにより前記識別を検証するステップを更に含む、請求項 2 に記載の方法。

20

【請求項 4】

前記未識別録音に関連付けられた前記テキストが前記録音のファイル名を含む、請求項 3 に記載の方法。

【請求項 5】

前記未識別録音に関連付けられた前記テキストが前記録音の ID 3 タグを含む、請求項 3 に記載の方法。

【請求項 6】

前記少なくとも 1 つのアルゴリズムにより決定された音紋が、前記未識別録音の少なくとも 1 つの部分の音声およびビデオ情報の少なくとも一方から抽出される、請求項 1 に記載の方法。

30

【請求項 7】

前記少なくとも 1 つのアルゴリズムにより決定された音紋が少なくとも 2 つの音紋を含み、
前記識別ステップが各々の信用度で前記単一基準録音を識別するために前記少なくとも 2 つの音紋の各々を必要とする、請求項 1 に記載の方法。

【請求項 8】

前記未識別録音の抽出された長さを前記単一基準録音の保存された長さと比較することにより前記識別ステップを検証するステップを更に含む、請求項 1 に記載の方法。

【請求項 9】

複数のユーザーからの複数の未識別録音について前記抽出、決定、識別の各ステップを繰り返すステップと、
前記未識別録音の識別の失敗を監視するステップと、
誤りのパターンから前記識別データベースに含まれ得る見込み誤りを検出するステップと、
を更に含む、請求項 8 に記載の方法。

40

【請求項 10】

前記監視ステップが前記ユーザーから、前記識別ステップが不正確なことを示す情報を受信するステップを含む、請求項 9 に記載の方法。

【請求項 11】

前記監視ステップは、前記検証ステップが成功しなかった場合、前記少なくとも 1 つのアルゴリズムにより決定された音紋と前記単一基準録音の識別子とを保存するステップを含

50

み、
単一の識別子に対して実質的に異なる音紋が保存されている場合に、前記識別データベースに前記見込み誤りがあることを示すステップを更に含む、請求項 9 に記載の方法。

【請求項 12】

前記少なくとも 1 つのアルゴリズムにより決定された音紋が前記基準録音の 1 つに合致するが、前記未識別録音には前記合致した 1 つの基準録音のものとは異なる ID3 タグ情報が関連付けられている場合に、前記見込み誤りを示すステップを更に含む、請求項 9 に記載の方法。

【請求項 13】

前記未識別録音から抽出された情報に基づいて前記見込み誤りを訂正するステップを更に含む、請求項 9 に記載の方法。

【請求項 14】

前記少なくとも 1 つのアルゴリズムにより決定された音紋が前記基準録音のうちの一つと実質的に類似するが、前記未識別録音からは実質的に異なる情報が抽出された場合に、前記識別データベースに見込み誤りを示すステップを更に含む、請求項 8 に記載の方法。

【請求項 15】

表示およびユーザーによる操作を目的として、前記未識別録音内に埋められたデータを補足すべく前記未識別録音に関連するデータを補足データベースから配信するステップを更に含む、請求項 1 に記載の方法。

【請求項 16】

録音を識別する方法であって、
複数のアルゴリズムを用いて未識別録音の少なくとも 1 つの部分から複数の音紋を抽出するステップと、
前記未識別録音から抽出された音紋のうちの少なくとも 2 つと基準録音に対応して生成された音紋を含む少なくとも 1 つのデータベースとを用いて前記未識別録音の見込み識別を決定するステップと、
各音紋に基づく前記見込み識別が単一の基準録音を各々の信用度で識別した場合に前記未識別録音を識別するステップと、を含む方法。

【請求項 17】

各音紋が前記未識別録音の少なくとも 1 つの部分の音声およびビデオ情報の少なくとも一方から抽出される、請求項 16 に記載の方法。

【請求項 18】

前記未識別録音の長さを前記単一基準録音の保存された長さと比較することにより前記識別ステップを検証するステップを更に含む、請求項 16 に記載の方法。

【請求項 19】

前記抽出ステップがユーザー所有のクライアント装置により実行され、
前記決定、識別、および検証の各ステップが識別サービス・プロバイダの制御下にある少なくとも 1 つのサーバーにより実行され、
前記検証ステップが成功した後で前記少なくとも 1 つのサーバーから前記クライアント装置へ前記単一基準録音に関連付けられた一意な識別子を送信するステップと、
前記一意識別子を前記クライアント装置内の前記未識別録音に関連付けるステップと、を更に含む、請求項 18 に記載の方法。

【請求項 20】

前記一意識別子を、前記少なくとも 1 つのデータベースに保存された識別子に関する許可リストと比較するステップと、
前記許可リストに前記一意識別子に合致するものがある場合、前記録音が共有できることを示すステップと、を更に含む、請求項 19 に記載の方法。

【請求項 21】

前記一意識別子を前記少なくとも 1 つのデータベースに保存された識別子のブロック・リストと比較するステップと、

10

20

30

40

50

前記ブロック・リストに前記一意識別子に合致するものがある場合、前記録音が共有できないことを示すステップと、を更に含む、請求項 19 に記載の方法。

【請求項 22】

未識別録音を識別すべく用いられるデータベースに保存されている基準情報を取得する方法であって、

ユーザー所有の録音を識別するための前記データベースの前記ユーザーが所有する録音に関連付けられた非波形データを取得するステップと、

前記録音の少なくとも 1 つの部分から少なくとも 1 つの音紋を抽出するステップと、

前記データベース内で前記非波形データと合致するものが検出された場合、前記少なくとも 1 つの音紋を前記録音の識別情報として保存するステップと、を含む方法。

10

【請求項 23】

前記非波形データが前記録音の長さを示す、請求項 22 に記載の方法。

【請求項 24】

前記録音がリムーバブルメディアに永久保存されており、前記非波形データが前記録音の TOC (Table of Contents) データから得られる、請求項 23 に記載の方法。

【請求項 25】

前記非波形データが前記録音に関連付けられたテキストを含む、請求項 22 に記載の方法。

【請求項 26】

前記非波形データが ID3 タグを含む、請求項 25 に記載の方法。

20

【請求項 27】

前記 ID3 タグが符号化情報を含み、

前記データベースが前記識別サービス・プロバイダにより管理されており、前記符号化情報が前記識別サービス・プロバイダの制御下で生成され、

前記データベースへの識別情報の保存ステップに先立って前記符号化情報を復号化することにより前記非波形データを検証するステップを更に含む、請求項 26 に記載の方法。

【請求項 28】

前記非波形データが電子透かしを含む、請求項 25 に記載の方法。

【請求項 29】

前記非波形データが音源媒体の種類に関する媒体情報を含む、請求項 25 に記載の方法。

30

【請求項 30】

前記媒体情報が前記音源媒体の種類を CD-R として識別する、請求項 29 に記載の方法。

【請求項 31】

前記媒体情報が前記音源媒体の種類を CD-DA として識別する、請求項 29 に記載の方法。

【請求項 32】

前記媒体情報が前記音源媒体の種類をデジタル・ファイルとして識別する、請求項 29 に記載の方法。

【請求項 33】

前記媒体情報が前記音源媒体の種類を DVD (デジタル多用途ディスク) として識別する、請求項 29 に記載の方法。

40

【請求項 34】

前記テキストが前記録音のファイル名を含む、請求項 25 に記載の方法。

【請求項 35】

前記テキストが前記録音のタイトルを含む、請求項 25 に記載の方法。

【請求項 36】

前記テキストが前記録音の制作に参画したアーティスト名を含む、請求項 25 に記載の方法。

【請求項 37】

50

前記テキストが前記録音に関連付けられたアルバム名を含む、請求項 25 に記載の方法。

【請求項 33】

前記取得および抽出ステップが、複数のユーザーが所有するクライアント装置により前記録音の異なる複製について実行され、異なるユーザーが前記録音から異なる音紋を抽出する、請求項 22 に記載の方法。

【請求項 39】

識別サービス・プロバイダの制御下にある少なくとも 1 つのサーバー上でデータベースを管理するステップと、

前記少なくとも 1 つのサーバーから前記クライアント装置へ、各クライアント装置が異なる音紋を抽出する旨の抽出命令を送信するステップと、を更に含む、請求項 38 に記載の方法。

【請求項 40】

前記クライアント装置から前記少なくとも 1 つのサーバーへ前記非波形データを送信するステップと、

前記非波形データに基づいて、前記少なくとも 1 つのサーバーによる前記クライアント装置への送信ステップ用に前記抽出命令を選択するステップと、を更に含む、請求項 39 に記載の方法。

【請求項 41】

前記録音に関する前記非波形データの受信頻度に少なくとも部分的に基づいて前記抽出命令を更新するステップを更に含む、請求項 40 に記載の方法。

【請求項 42】

前記抽出命令の前記選択ステップが、前記抽出命令を受信する前記クライアント装置の種類に少なくとも部分的に基づいている、請求項 40 に記載の方法。

【請求項 43】

前記抽出命令の前記選択ステップが、前記抽出命令を受信する前記クライアント装置の地理的位置に少なくとも部分的に基づいている、請求項 40 に記載の方法。

【請求項 44】

前記抽出命令の前記選択ステップが、前記抽出命令を受信する前記クライアント装置で実行されるソフトウェアに少なくとも部分的に基づいている、請求項 40 に記載の方法。

【請求項 45】

前記識別情報を提供したユーザーの数に少なくとも部分的に基づいて前記抽出命令を更新するステップを更に含む、請求項 40 に記載の方法。

【請求項 46】

前記抽出命令の前記選択ステップが、前記録音の複製の品質に少なくとも部分的に基づいている、請求項 40 に記載の方法。

【請求項 47】

前記抽出命令により指定される時間に、前記少なくとも 1 つの音紋を前記クライアント装置から前記少なくとも 1 つのサーバーへ送信するステップを更に含む、請求項 40 に記載の方法。

【請求項 48】

指定された数の音紋が前記送信ステップ用に準備できるまで、前記少なくとも 1 つの音紋を前記クライアント装置に保存するステップを更に含む、請求項 47 に記載の方法。

【請求項 49】

前記少なくとも 1 つのサーバーとの通信路が利用できる場合、前記少なくとも 1 つの音紋の前記送信ステップが行われる、請求項 47 に記載の方法。

【請求項 50】

クライアント装置の 1 つからアクセスされた第一の録音に対する前記少なくとも 1 つの音紋の前記送信ステップが、前記 1 つのクライアント装置からアクセスされた第二の録音に対する非波形データの前記送信ステップと共に行われる、請求項 47 に記載の方法。

【請求項 51】

前記録音がリムーバブルメディアに永久保存されて、前記クライアント装置が前記録音から少なくとも1つの符号化されたファイルを生成し、
前記録音の符号化に先立って前記送信ステップが前記少なくとも1つの音紋を送信する、請求項47に記載の方法。

【請求項52】

前記録音がリムーバブルメディアに永久保存されて、前記クライアント装置が前記録音から少なくとも1つの符号化されたファイルを生成し、
前記リムーバブルメディアの1つのトラックを符号化した後で、前記送信ステップが前記少なくとも1つの音紋を送信する、請求項47に記載の方法。

【請求項53】

前記録音がリムーバブルメディアに永久保存されて、前記クライアント装置が前記録音から少なくとも1つの符号化されたファイルを生成し、
前記リムーバブルメディアの符号化が完了したとの通知を受信した後で、前記送信ステップが少なくとも1つの音紋を送信する、請求項47に記載の方法。

【請求項54】

前記データベースが楽曲録音に関する識別情報を含む、請求項22に記載の方法。

【請求項55】

前記データベースがビデオ録画に関する識別情報を含む、請求項22に記載の方法。

【請求項56】

前記少なくとも1つの音紋の品質を検出するステップと、
前記少なくとも1つの音紋を用いて、前記録音の別の複製を識別するステップと、
前記録音の別の複製がより高品質な音紋を生成する場合に、前記少なくとも1つの音紋を前記より高品質な音紋と置換するステップと、を更に含む、請求項22に記載の方法。

【請求項57】

前記品質の検出ステップが、前記録音に使用した符号化技術に基づいている、請求項56に記載の方法。

【請求項58】

前記品質の検出ステップが、前記録音の保存に用いた媒体の種類に基づいている、請求項56に記載の方法。

【請求項59】

前記品質の検出ステップが、前記録音にアクセスするユーザー装置の誤り訂正性能に基づいている、請求項56に記載の方法。

【請求項60】

前記ユーザー装置がソフトウェア誤り訂正ではなくハードウェア誤り訂正を用いた場合に、前記品質の検出ステップがより高い品質を与える、請求項59に記載の方法。

【請求項61】

前記品質の検出ステップが、前記音紋の前記抽出ステップの間に検出された誤りの数に基づいている、請求項56に記載の方法。

【請求項62】

前記取得および抽出ステップが、前記録音の異なる複製について複数のユーザーが所有するクライアント装置により実行され、
前記ユーザーの1人から取得された前記少なくとも1つの音紋を、少なくとも1人の別のユーザーから抽出された前記少なくとも1つの音紋と比較するステップと、
前記比較ステップに基づいて前記データベースの前記少なくとも1つの音紋を更新するステップと、を更に含む、請求項22に記載の方法。

【請求項63】

前記比較ステップが、異なるユーザーからの音紋が所定の相関を有すると判定した後で、前記更新ステップが実行される、請求項62に記載の方法。

【請求項64】

前記更新ステップが前記データベースに保存すべく異なるユーザーからの音紋を組み合わせ

10

20

30

40

50

せる、請求項 6 2 に記載の方法。

【請求項 6 5】

録音を識別するシステムであって、

ユーザー所有の媒体に保存された未識別録音に関する情報と少なくとも 1 つのアルゴリズムにより決定された音紋とを、前記未識別録音の少なくとも 1 つの部分から抽出する抽出ユニットと、

前記抽出ユニットに接続されており、前記未識別録音から抽出された情報の少なくとも一部と、基準録音の対応する情報を含む識別データベースを用いて、前記未識別録音の見込み識別を行なうとともに、前記少なくとも 1 つのアルゴリズムにより決定された音紋と組み合わされた前記情報の少なくとも一部の各々に基づく前記見込み識別が、単一の基準録音を各々の信用度で識別した場合に前記未識別録音を識別する識別ユニットと、を含むシステム。

10

【請求項 6 6】

録音を識別するシステムであって、

複数のアルゴリズムを用いて未識別録音の少なくとも 1 つの部分から複数の音紋を抽出する抽出ユニットと、

前記抽出ユニットに接続されており、前記未識別録音から抽出された少なくとも 2 つの音紋と基準録音に対応して生成された音紋を含む少なくとも 1 つのデータベースとを用いて前記未識別録音の見込み識別を行なうとともに、各音紋に基づく前記見込み識別が単一の基準録音を各々の信用度で識別した場合に前記未識別録音を識別する識別ユニットと、を含むシステム。

20

【請求項 6 7】

未識別録音を識別するために用いられるデータベースに保存された基準情報を取得するシステムであって、

ユーザー所有の録音を識別するための前記データベースの前記ユーザーが所有する録音に関連付けられた非波形データを取得する受信ユニットと、

前記録音の少なくとも 1 つの部分から少なくとも 1 つの音紋を抽出する抽出ユニットと、前記受信ユニットおよび前記抽出ユニットに接続されており、前記データベース内で前記非波形データに対する合致が検出された場合、前記録音の識別情報として前記少なくとも 1 つの音紋を保存する記憶ユニットと、を含むシステム。

30

【発明の詳細な説明】

【技術分野】

【0001】

本発明はコンテンツから録音／録画を識別すること、特に音紋識別と録音／録画に関するその他の情報を組み合わせて、始めは最も廉価な識別方式を用いながら必要に応じてより複雑な方式を重ねていくことにより識別の信頼性を向上させて信頼性の高い識別を効率的に実現させることを指向する。

【背景技術】

【0002】

音声（およびビデオ）録音／録画の識別には各種の用途がある。多くの用途は、録音／録画された作品の再生および演奏に対する補償や権利保有者による統制に関する。20 世紀末から 21 世紀の初めにかけて開発されたナップスター等のファイル共有ソフトウェアやその他同様のサービスにより、この種のシステムを利用することの重要性が増した。正確な識別を行うことに対するニーズは数年来高かったが、どのシステムもこのニーズに応えることができなかった。

40

【0003】

録音／録画識別の別の用途は、録音を聞く（または録画を見る）際に、ユーザーに対し付加価値を提供することである。一例がカリフォルニア州バークレーのグレースノート社（Gracenote, Inc. Berkeley, California）から提供されている CDDB 楽曲識別サービスであり、コンパクトディスク（CD）を識別して、識別

50

されたCDに関する情報、例えばアルバム名、アーティスト、トラック名、およびインターネット上の関連コンテンツ（アルバムカバー、アーティストおよびファンのウェブサイト等）へのアクセスを提供する。CDDBサービスはコンパクトディスクを識別することに関しては有効であるものの、CDやDVD等のリムーバブルディスクに保存されていないファイルの識別に用いる場合にはいくつかの短所がある。

【発明の開示】

【発明が解決しようとする課題】

【0004】

全ての音声音紋技術には「盲点」、すなわちその技術を用いるシステムが音声の類似や相違を誤って識別する箇所がある。ただ一つの音紋技術に依存している限り、盲点に遭遇した際にその技術により得られる解が不正確なものとなる。

10

【0005】

グレースノート社製CDDBシステムは、さらに音声ファイルをMP3その他のフォーマットへデジタル的に符号化するアプリケーションに使用されることで有名である。これらの符号化アプリケーションはグレースノート社のCDDBサービスを利用して、符号化の対象とされたコンパクトディスクを識別し、タイトルおよびIDタグに正しいメタデータを書き込む。グレースノート社のCDDBサービスは、各トラックに対して一意なID（TUID）を返し、この種のIDをMP3ファイル用のID3V2タグに挿入する。TUIDはハッシュ化済みかつ権利保護されており、グレースノート社製システムでのみ読み込み可能である。しかしID3V2タグは簡単に編集することができ、あるファイルのTUIDを別のファイルのID3V2タグに保存することが可能である。従って、TUID単独ではファイルの音声コンテンツに対して信頼性の高い識別子とはいえない。

20

【0006】

グレースノート社製CDDBシステムはまた、クライアント装置により抽出されたテキストをトラック、アーティスト、およびアルバム名が格納されたメタデータ・データベースに対し一致検索することにより、ファイル名、ファイル・パス、IDタグ（タイトル）等からデジタル音声ファイルを識別するのに利用されるテキスト・マッチング機能も提供する。このテキスト・マッチングは識別力を向上させるため各録音に対してユーザが入力した内容も関連付けて利用するが、録音がコンパクト・ディスクから切り離されて、いずれかのフォーマットでファイルに保存されてしまうと、録音の音声コンテンツとテキストが合致することを確認する方法はなかった。

30

【課題を解決するための手段】

【0007】

本発明の一態様では、リソースの使用を最少限に抑えつつ、録音の識別能力を最大化する。

【0008】

本発明の別の態様では、複数の識別方法を用い、音声音紋等のリソースを多く消費する方法は必要な場合のみに用いるようにする。

【0009】

本発明の更なる態様では、未識別データの処理を最少限に抑える。

40

【0010】

本発明の更に別の態様では、所望の信用度に達するまで、最も廉価な識別技術に、徐々により高価な識別技術を付加して用いる。

【0011】

本発明のまた更なる態様では、未識別録音が関連付けられたテキストと、識別済み録音に関連付けられたテキストを比較することにより、コンテンツに基づく録音の識別を検証する。

【0012】

本発明のまた更に別の態様では、信頼性を向上させるべく、異なる音源からの録音識別方法を利用する。

50

【0013】

本発明のなお更なる態様では、ファジー・トラック長分析を用いて、コンテンツに基づく録音の識別を検証する。

【0014】

本発明の更に別の態様では、基準データベースで用いられる、録音を識別用の識別データを自動的に抽出する。

【0015】

本発明のなお更なる態様では、未識別録音を周期的にシステムへ再投入して、最近追加されたデータまたは最近改良された技術により識別可能か否かを判定する。

【0016】

上述の態様は次のような録音識別方法により実現可能である。同方法は、ユーザー所有の媒体に保存された未識別録音に関する情報と、当該未識別録音の少なくとも1つの部分から少なくとも1つのアルゴリズムにより決定された音紋とを抽出するステップ、当該未識別録音から抽出された情報の少なくとも一部と基準録音の対応する情報を含む識別データベースを用いて当該未識別録音の見込み識別を行なうステップ、および当該情報の少なくとも一部の各々とアルゴリズム的に決定された少なくとも1つの音紋を組み合わせたものに基づく前記見込み識別が単一の基準録音をそれぞれの信用度で識別する場合に当該未識別録音を識別するステップによる。当該未識別録音の少なくとも1つの部分は、音声、ビデオ、またはその両方を含んでいてよい。

【0017】

データベースは好適には、識別サービス・プロバイダの制御下にあるサーバーだけが識別できる一意な識別子を提供する識別サービス・プロバイダにより管理される。一意識別子は、一旦識別された録音に関係付けられる。続いて、一意識別子を用いて録音の複製が識別されて、処理速度が極めて高速になる。一意識別子はオプション的に、高速RAMまたは特別に索引付けられたデータベース・テーブルにキャッシュされる。

【0018】

未識別録音に非波形データが利用できない場合、当該未識別録音は、複数のアルゴリズムを用いてその未識別録音の少なくとも1つの部分から音紋を抽出するステップと、当該未識別録音から抽出された少なくとも2つの音紋と基準録音に対応して生成された音紋を含む少なくとも1つのデータベースとを用いて、当該未識別録音の見込み識別を行なうステップと、各々の音紋に基づく見込み識別が各々の信用度で単一の基準録音を識別した場合に、当該未識別録音を識別するステップにより識別されることが好適である。

【0019】

好適には、ユーザー所有の録音の識別に用いられ、音紋情報を含まない既存のデータベースが、データベースのユーザーが所有する録音に関連付けられた非波形データを取得するステップと、録音の少なくとも1つの部分から少なくとも1つの音紋を抽出するステップと、データベースにて非波形データと合致するデータが検出された場合に当該少なくとも1つの音紋を当該録音の識別情報として保存するステップにより拡張される。一例として、ユーザー所有のオーディオCDからデジタル楽曲ファイルを符号化する処理の間に識別システムを用いてオーディオCDを識別し、固有IDシステムを用いて符号化処理の間に抽出された音紋をオーディオCDに直接関連付けることができる。

【0020】

音紋または一意識別子のいずれかを用いる録音の識別は、好適には識別用データベースに保存されている他の情報、例えば録音の長さや録音内部に埋め込まれた数値識別子等により検証される。検証に合格しないか、または識別に使用する情報のいくつかには合致するものの全てには合致しない録音に関する情報は、誤りの原因を後で分析すべく保存することができる。上述のように音紋が得られた場合、音紋の取得に誤りがあった可能性がある。従って、マッチング失敗用に保存されている情報のパターン識別に基づいて、誤りをオペレータに対して出力することも、あるいはデータベースに保存されている情報をシステムが修正することもできる。例えば、大多数の合致した音紋が保存されていても、それ

10

20

30

40

50

外の情報は一貫してそれらに合致しない場合には、音紋データベースの誤りをオペレータに通知する必要があるかもしれない。

【0021】

本発明は、ユーザー所有の媒体に保存された未識別録音に関する情報と少なくとも1つのアルゴリズムにより決定された音紋を、当該未識別録音の少なくとも1つの部分から抽出する抽出ユニットと、当該抽出ユニットに接続されており、当該未識別録音から抽出された情報の少なくとも一部と基準録音の対応する情報を含む識別データベースとを用いて当該未識別録音の見込み識別を行ない、当該情報の少なくとも一部の各々とアルゴリズム的に決定された少なくとも1つの音紋との組み合わせに基づく当該見込み識別が、単一の基準録音を各々の信用度で識別した場合に当該未識別録音を識別する識別ユニットを含む録音識別システムを含む。

10

【0022】

本発明はまた、複数のアルゴリズムを用いて未識別録音の少なくとも1つの部分から音紋を抽出する抽出ユニットと、前記抽出ユニットに接続されており、当該未識別録音から抽出された少なくとも2つの音紋と基準録音に対応して生成された音紋の少なくとも1つのデータベースとを用いて、当該未識別録音の見込み識別を行ない、各音紋に基づく当該見込み識別が、単一の基準録音を各々の信用度で識別した場合に当該未識別録音を識別する識別ユニットを含む録音識別システムを含む。

【0023】

上述のシステムのいずれにおいても、抽出ユニットは通常、インターネット等のネットワークを介して識別ユニットとしての少なくとも1つのサーバーに接続されたクライアント装置である。クライアント装置は、録音にアクセスするドライブを備えたパーソナル・コンピュータ、ネットワーク接続を有する家電製品、または未識別録音のある場所から別の場所へ送信するサーバー・コンピュータであってよい。さらに、データベースの一部はローカルに利用可能であって、抽出ユニットおよび識別ユニットが同一の装置に配置されてコンポーネントを共有することができる。

20

【0024】

本発明はまた、未識別録音を識別するために用いられるデータベースに保存された基準情報を取得するシステムを含む。当該システムは、ユーザー所有の録音を識別するためのデータベースのユーザーが所有する録音に関連付けられた非波形データを取得する受信ユニットと、録音の少なくとも1つの部分から少なくとも1つの音紋を抽出する抽出ユニットと、前記受信ユニットおよび前記抽出ユニットに接続されており、データベース内で非波形データに対する合致が検出された場合に録音の識別情報として少なくとも1つの音紋を保存する記憶ユニットを含む。

30

【0025】

これらおよび引き続き開示される他の態様や利点は、本明細書および請求項に記載する詳細な構成および動作に含まれており、本明細書の一部をなす添付図面を参照するが、その際同一部材には同一番号が付与されている。

【発明を実施するための最良の形態】

【0026】

本発明によれば、図1に示されるようなシステムにおいて識別に使用する装置がセットになってコンポーネントの組が提供され、複数の方法を利用して音声（およびビデオ）ファイルの分析および識別を行うことを容易にする。好適には、録音識別子およびテキスト・データを含むデータベース90をデジタル音声をテキスト情報に基づき識別する方法と音紋を用いて音声を識別する方法と組み合わせる。好適には、データベース90内のテキスト・データはユーザーにより与えられ、ユーザーから入力されたスペリングのバリエーションを含んでいる。そのようなデータベースの一つとして、グレースノート社のCDDB音楽識別サービスが利用可能である。

40

【0027】

図1に示すように、録音100は、ハードディスクまたはコンパクトディスクからデジタ

50

ル音声ファイルを読み込む等の任意の従来方法によりクライアント装置 110 からアクセスされる。そして、録音 100 および付随する情報（メタデータ）から情報が抽出される。以下に詳述するように、録音 100 から抽出される情報は音紋である。メタデータから抽出された情報は、CDトラックの（TOCから得られる）トラック長である録音の演奏時間、録音がMP3ファイルにある場合はファイル名およびID3タグ、および録音がCD上にある場合は目次（TOC）データを含む。録音を含んでいるファイルが本発明に従って動作するクライアント装置により生成された場合、ID3ファイルから一意なIDが抽出されることになるが、最初はその情報が利用できないものと仮定する。

【0028】

一実施形態の例において、抽出された情報はクライアント 110 からサーバー 120 へ送られ、録音 100 から抽出された情報の少なくとも一部および基準録音に対応して生成された音紋を含むデータベース 130 を用いて未識別録音の見込み識別を行なう。テキスト情報または一意なID（固有ID）が抽出されれば、一致検索が行われる。テキスト情報または一意なIDを用いて合致が検出された場合、少なくとも1つのアルゴリズムにより決定された音紋が、合致記録に含まれる音紋と比較されて、録音 100 から抽出された情報に合致し、合致する情報の各項目について信用度が十分な基準を満たす単一の基準録音が存在するか否かを判定する。テキスト情報および一意なIDに基づいた合致が検出できない場合、録音 100 から抽出された音紋のうちの少なくとも2つを用いて単一の基準録音に識別できるかが試みられる。いずれかの方法を用いて単一の基準録音が特定された場合、好適には最終検証ステップとして録音 100 の演奏時間が単一の基準録音の演奏時間と比較される。

【0029】

好適には、関連メタデータを用いて音紋識別により得られた合致を検証する。あらゆる識別システムと同様に、音紋は誤った結果をもたらす場合がある。検証用コンポーネントが無ければ、この種の誤りはシステム全体に伝播して、大多数のユーザーに誤ったデータを返す恐れがある。トラック長の比較等によって検証を行うことにより、システムが潜在的な誤りを捕捉し、その誤りに検証用のフラグを立てておくことが可能になる。

【0030】

本発明によるシステムは好適には、各種の識別方法およびその重要度の重み付けを可能にする、独自仕様の結果報告用および柔軟な管理用インターフェース 130 を含む。特定の識別方法毎に合致成功率を分析することにより、管理者は識別確率を最大化すべく各構成要素の識別基準を変えることができる。本発明によるシステムは好適には、検索結果（140）を補助すべく、グレースノート社のデータサービス部門を介してCDDBデータベースを利用している2800万人以上のユーザーによる利用データを反映させる。

【0031】

本発明によるシステムは柔軟性を有するため、異なる環境下で録音を識別する場合、異なる構成にすることができる。例えば、ストリーミング音声監視するアプリケーションは、ピア・ツー・ピア・システムのファイルを識別するものやアナログ入力を識別するものとは全く異なるシステムおよび構造を必要とする。しかし、本発明はこれらの各状況下で録音を識別できるように構成することができる。

【0032】

本発明によるシステムはリソース消費量を最少限に抑えつつ、識別機能を最大化する。複数の識別方法を利用して、音声音紋等のリソースを多く消費する方法は必要なときのみを用いる。複数の音声音紋技術を利用することにより、データコリジョンが減り、既存の音声音紋技術のあらゆる「盲点」がカバーされる。単一の音紋技術を利用する音紋システムに見られる「盲点」は、複数の異なる音紋技術を用いることにより回避される。これはまた、特定の用途に用いる際の微調整を可能にする。

【0033】

好適には、音紋は図2に示す方法を用いる複数音紋識別サービスを利用して取得される。これにより、システムは各種の録音を正確に識別することができる。

【0034】

図2に示すように、未識別（未知の）録音100が音紋を抽出するクライアント装置110からアクセスされた場合、可能ならば識別システム210により従来のTOC／ファイル識別が実行されて、結果220がクライアント装置110に返される。TUIDが検出された場合、結果220には、マスター・メタデータ・データベース（図2に示さず）の内容へのポインタとなる一意識別子（TUID）が含まれる。録音100はまた、少なくとも1つ、好適には複数の、それぞれアルゴリズムが異なる音紋抽出システムを用いて音紋抽出部230によって処理され、音紋／ID送信キャッシュ240に保存される。以下に詳述するように、音紋アップローダー250が音紋をいつ音紋識別サーバー120へ送信すべきかに関する指示が受信される。

10

【0035】

音紋識別サーバー120において、音紋アップローダー250により送信された音紋はまずは音紋受信キャッシュ260に保存される。次いで音紋は、録音の音紋と異なるエンドユーザーにより抽出されアップロードされた音紋との相関性を調べるためのアルゴリズム比較器を用いて音紋検証270を受ける。音紋同士がかなり類似していることがわかった場合に検証が行われる。検証は上記以外の方法によっても行うことができるが、上記方法は不良データの排除に役立つプロセスの一例である。

【0036】

本実施例において、正当かつ関連があると判定された音紋は、音紋縫合部280へ送られる。例えば、音紋が録音を30秒に区切った複数の区間から取られている場合、音紋は連続的な音紋ストリームに加工される。これにより、録音区間の識別が単純化できる。結果的に得られた音紋は、既存データベース90（図1）に関連付けられた音紋データベース290に保存される。

20

【0037】

CDDBデータベースは一部がユーザーから与えられた内容により構築されており、2002年半ば時点には1200万トラックおよび900,000アルバムを超えるメタデータ・データベースが構築されている。このデータベースは、基本メタデータ（アーティスト、アルバム、およびトラック名）と拡張データ（ジャンル、ラベル等）の両方を含む。

【0038】

図1に示すシステムを用いて、波形データベースの構築に同様の分散型収集方法を利用することができる。録音100が生音声波形である場合、CDがMP3ファイル等の別のフォーマットに符号化されるときに、クライアント装置110は、データベース90のユーザーにより所有される録音100に関連付けられた非波形データを取得し、録音の少なくとも1つの部分から音紋を抽出する抽出アルゴリズムを実行する。次いで音紋は、好適にはCDのTOCから得られるような一意なIDと共にサーバー120へ送信される。一意なIDが利用できる場合、すなわち、データベースで非波形データに対する合致が検出された場合、サーバー120はデータベース90の適切なメタデータと、ユーザー所有の録音を識別するために得られた音紋を、データベース90を用いてCDを識別するのと同レベルの精度で関連付けることができる。この方法で動的に集められた音紋は以下に詳述するように、録音の識別情報として少なくとも1つの音紋を保存するに先立って、認証済みクライアントからの音紋を蓄積する音紋収集サーバー（図1に示さず）へ送信されてもよい。

30

40

【0039】

複数の音紋を収集する抽出部はまた、ラウドアイ（Loudeye）やミューズ（Muse）等の市販エンコーダに含まれる静的波形の組にも適用される。このアプローチの課題は、音紋を適切なメタデータに関連付けることである。上述の方法により、音声音紋を親レコードに論理的に関連付けて、オリジナル音源へ逆向きに関連付けることができる。好適な実施例においては、一意なIDにより同じ曲のライブとスタジオ・バージョンが区別され、同時にそれらの録音を同一アーティストおよび各々のアルバムにリンクすることができる。

50

【0040】

好適には、サーバー120は情報を、一意なIDに関連付けられた並列な記録の集合として保存する。クライアント110がサーバー120に対して媒体(CD、デジタル音声ファイル、ビデオファイル)の識別を依頼すると、サーバー120もまた、この特定のCDについて音紋をどのように収集すべきかに関する記録を返す。これは、収集指示レコード(GIR)と呼ばれる。GIRは、一連の指示を含んでもよい。その後には、リモート音紋収集コードが続く。このレコードは、負荷の少ない時間に予め計算しておくことも、または識別時に動的に計算することもできる。

【0041】

サーバー120は、CDの人気に関して既知の情報を利用して、収集に関する意思決定を行う。音紋を取得する機会を逃がさないよう、稀少CDに関するあらゆる項目が収集される(たとえユーザーには多少負担であろうとも)。非常に人気のあるCDの場合は正反対の状況である。その負担は多くのユーザーに分散されるため、彼らは音紋収集を手伝わされていることに気付くもしないだろう。

【0042】

GIRを構築するための規則および手順は、手動でも自動化されていてもよく、時間経過に伴ない変更されてもよい。これらはまた、特定のユーザー、アプリケーション、または地理的位置に個別に適用することができる。

【0043】

一実施例において、サーバーは以前に収集された音紋を削除するようにGIRを修正して動的に音紋を収集する。GIRの更新頻度は即時から日、週、または月の範囲で変動する。GIRに含まれる指示としては次のようなものがある。

一収集されるトラックと区間、およびそれらの優先度のリスト。

一音紋ジェネレータが使用するアルゴリズム。

一音紋ジェネレータに音紋を処理する方法を指示するパラメータ。例えば、音声サンプルの周波数

処理すべき音紋の周波数帯域

音紋の解像度

音声の要求品質

一音紋収集をいつ実行するか、例えば、

トラック符号化の前

トラック符号化の後

トラック符号化と並行して

一音紋をキャッシュに書き込んで、それをサーバーへ返送する時期に関する指示。例えば、

、

トラック符号化の前

トラック符号化の後

CD全体が符号化された後

サーバーへの返送用通信チャネルがビジーでないとき

次のCDを調べるとき

音紋群が送信準備状態にあるとき

一コンピュータに過負荷をかけないようにCPUパワーを処理に投入する指示

好適には、システムは動作の間、音紋の品質を向上させようと試みる。音源信号の品質や音紋抽出に使用するパラメータは、より高品質なものが利用できる場合改善されてゆき、その結果、サーバー120がどの音紋を収集すべきかを決定するために使う品質行列が複雑になる。音源品質の例を以下に挙げる。好適には、データベース90または音紋収集サーバーが管理する類似データベースが、データベースに格納されている音紋の音源品質を保存しているため、より高品質な音源からの音紋が利用できる場合、音紋は入れ替えられる。

【0044】

10

20

30

40

50

【表 1】

音源品質表

名前	ビットレート	圧縮	誤り訂正	品質指数
CD_Audio_HEC	44100kbps	無し	ハードウェア	1
CD_Audio_SEC	44100kbps	無し	ソフトウェア	2
CD_Audio	44100kbps	無し	無し	3
CDR_Audio	44100kbps	無し	無し	4
CDR_Made_From_MP3	44100kbps	mp3	無し	5
MP3_File	160kbps	mp3	無し	6

動的に収集された音紋に、品質の検証に役立つ情報を含めることができる。媒体からの読み込む際の誤り率等の情報が音紋収集機能へ送信される。本システムは、音源媒体からの誤り率の高い音紋を拒絶することができる。

【0045】

上述のように、音紋を直ちに保存するのではなく、録音の複数の音紋が、音紋収集サーバーにより収集されデータベースに追加されてもよい。これらの音紋をアルゴリズムを用いて比較してそれらの相関性が判定される。十分な相関がない場合、十分な相関が得られるまでさらに音紋を収集することができ、音紋のうち1つまたは合成された音紋がデータベースに保存される。これにより、不良音紋がデータベースに格納されることを防止する。

【0046】

タイミングがわずかに変しても音紋の重なりが生じる恐れがあるため、区分化された音紋の縫合が必要な場合がある。アルゴリズム化された縫合により、より高品質の連続的な音紋が得られる。単純な縫合のやり方は、区分化された音紋を録音における出現順に追加するものである。複雑な縫合のやり方は、音紋の異なる特性を最小共通項目に合わせて尺度調整してから、録音における出現順にそれらを追加するものである。好適には、区分けされた音紋がジッターを含む場合、何らかの数学的適合判定が用いられるため、追加を行うことはデータストリームへの単純な追加よりもファジーな処理である。

【0047】

利用可能な音声音紋の一例が、「録音の自動識別」と題された米国特許出願（マクスウェル・ウェルズ（Maxwell Wells）ら、2002年7月22日出願）に記載されており、本明細書で引用している。しかし、任意の公知のアルゴリズムを用いて導かれた音紋取得技術はデジタル音声に限らず、ビデオ、テレビ番組（アナログ、デジタル共に）、およびDVDの用途に用いることができる。特定の用途で識別される媒体に適当な識別子および識別技術が用いられる。

【0048】

本発明は優れた柔軟性を有し、ピア・ツー・ピア環境におけるMP3識別、あるいは監視や報告目的での音声ストリームの識別等、広範な環境で利用可能である。複数の識別コンポーネントを使用する以外にソリューションは見つかっておらず、本発明が任意の音声（またはビデオ）識別アプリケーションのニーズに応えるべくカスタマイズ可能な唯一のソリューションである。

【0049】

ピア・ツー・ピア・アプリケーションに本発明を実施するための機能的な説明を、図3を参照しつつ以下に述べる。本実施形態において、音声ファイルは一般からアクセス可能に

する前に識別され、ファイルがシステムに受け入れ可能であれか否かが判定される。これは「フィルター・イン」として知られる処理である。

【0050】

クライアント装置110（図1）は、サーバー120（図1）へのアップロード時点で音声ファイルから情報310（図3）を抽出する。抽出された情報は好適には、一意なID、ID3タグ、ファイル名テキスト・データ、トラック演奏時間等の非波形データ、および音紋を含む。これらは、録音から抽出されて識別のためにサーバー120へ送信される。

【0051】

最初の一致検索320は、一意なIDが存在すればそれに対して実行される。グレースノート社のTRIIDを利用すれば、99.9%の精度で検索結果を返すことが可能になる。これはまた、リソース消費が最も少ない識別方法であって、非常に高速な識別速度を実現可能である。一致する一意IDが存在すれば、システムは検証段階へ移行する。一致する一意IDが存在しない場合、システムは次の識別方法330を用いて識別を試みる。

【0052】

本実施例において次に、900,000アルバムおよび1200万曲を超える楽曲を含むグレースノート社製CDDBサービス等のメタデータ・データベースを用いてテキスト情報に基づく識別が行われる。テキスト情報を用いた一致検索は、ファイル名、ファイル・パス、またはMP3ファイルのID3タグ内のテキスト等、利用可能なテキスト情報を用いて、識別に使用するデータの集合を提供する。許容できる一致検索結果が返された場合、システムは検証段階へ移行する。一致するものがないと判定される場合、システムは次の識別方法を利用して識別を試みる。

【0053】

次のステップは音紋を利用する識別であり、この場合は音声音紋を用いる。未識別録音からの音紋が基準録音用のデータベース90の音紋と、一度に1つずつ（または異なる音紋に別々のプロセッサを用いて並行して）比較される。各音紋技術は、一致検索結果および信用度を返す。一致検索により単一の基準録音に絞られ、その信用度が許容可能な場合、システムは検証段階へ移行する。一致するものがないと判定される場合、システムは対象アプリケーションに応じて、ユーザーに最も一致すると思われる結果の検証を求めたり、あるいは、「一致するものがない」との結果を返すことができる。

【0054】

検証は、あらゆる識別システムが成功するための鍵となる要素である。好適には、録音の演奏時間等の鍵となるファイル属性を用いて、未識別録音から抽出された演奏時間を、データベースに格納された単一の基準録音の演奏時間と比較することにより、あるファイルと、識別システムが判定したものとが同一であるか否かを検証する。

【0055】

好適には、発見的（ヒューリスティック）多数決アルゴリズム340を用いて一致検索の結果がシステムが判定するものと同じであるか否かを決定する。この自己監視により、システムに悪影響を与える不正確なデータをシステムが検索結果として返す可能性を減少させることができる。発見的法則は最適合致を行うべく、手動制御またはアルゴリズムにより制御される。これらの発見的法則はまた、どの識別技術をどのような順番で適用すべきかを決定するのに用いてもよい。

【0056】

各アプリケーションの管理者は、システムの各ステージ（またはコンポーネント）で必要とされる精度のレベルを決定することができ、従ってシステムを最適化するための明示的な統制を行なうことができる。例えば、集計合致度が90%であることが必要な場合、システム管理者は管理インターフェース130を用いて受容可能な結果のレベルを90%に調整して、その閾値が満たされない限り一致が成立しないようにできる。管理者はまた、コンポーネント毎に結果レベルを設定することができる。例えば、テキストの合致度のレベルを99%に設定し、音声音紋の合致度のレベルを85%に抑えて設定できる。

10

20

30

40

50

【0057】

識別成功との結果が返されたならば、一意なIDでファイルに再びタグが付与されて350、システム全体にわたって正確なIDを持つファイルとして生息する。その結果、その後のファイル識別でのリソース消費が最少限で済む。

【0058】

ファイルに割り当てられた一意なID(TUID)は次いで、システムへの入力に許されたレコードレーベル、製作者、およびそれらのファイルのコンテンツ所有者が入力するタイトル/アーティストペア370により発生するTUIDのリスト370とマッチングされる(360)。一実施形態において、そのTUIDがデータベースに存在する場合はファイルの共有が許可されるが、TUIDがデータベースに存在しない場合にはファイルはブロックされる。別の実施形態では、TUIDがデータベースに存在する場合にファイルはブロックされる。これらの実施形態のいずれも、ファイルがユーザーからアクセスされたか、またはあるコンピュータから別のコンピュータへ送信されたかを識別するのに適用可能である。

10

【0059】

図4Aに示すように、本発明の実施形態は複数の関連データベースを使用する。マスター・メタデータ・データベース410は、タイトル、アーティスト/著者名、所有者名および日付の情報を含んでいる。関連データベースには、音紋データベース290(図2)を形成する音声音紋データベース430およびビデオ音紋データベース440を含んでいる。また、トラック長、TOCデータベース450、テキスト・データベース460、およびハッシュIDデータベース470並びに一意であることが確認された固有IDデータベース480も含まれる。

20

【0060】

図4Bに示すように、未識別(未知)録音100がクライアント装置110からアクセスされた場合、音紋540、550、メタデータ560、および存在すれば固有ID570を含む情報が抽出される。さらに、録音の演奏時間580が決定され、数値ハッシュ590が計算される。抽出された音紋は、音紋600、610と比較される。同様に、数値ハッシュ、テキスト、および一意ID(固有ID)に対する一致検索が実行される(620、630、640)。基準録音が特定されたならば、未識別録音100の演奏時間を基準録音の演奏時間と比較することにより検証が行なわれる。比較方法各々の信用度と共に検索された結果660~710が結果集計部730へ送られる。

30

【0061】

未識別録音100に合致する基準録音が検出されなかった場合(750)、抽出された情報540~590および結果は、周期的に識別サーバー120へ再投入するべく未識別一時保管箱760に保存される(図2、4B)。本実施例において、基準録音が特定されたものの信用度が低い場合(770)、例えば合致する可能性のある、見込み候補を1つ以上ユーザーに提示し、その応答としてユーザーの選択を受信する等の発見的方法790、または手動による見直し810を適用する等の識別後処理780が実行される。このようなユーザーの選択結果が発見的方法用データベース820に保存される。識別後処理780の結果単一の基準録音が識別されるか、または結果集計部730が識別結果の信用度が高いことを出力した場合(770)、ハッシュIDが生成810されてハッシュ・データベース480およびクライアント装置110へ送られる。このため、ファイルを作成するときにハッシュおよび一意ID(TUID)がID3タグに保存される。

40

【0062】

一実施形態においてシステムは、識別結果を向上させようと類似ファイルの識別を繰り返したときの誤り率に基づく学習を行う。また、結果に誤りがあることを示すユーザーから、手動で通知を受けることもできる。これにより、時間の経過とともに識別を連続的に検証することができる。例えば、本発明のシステムによりファイルを識別することができ、次いで時間の経過とともにシステムはそのファイルの識別に欠陥があると判定して、オペレータにその旨を通知する。別の実施形態において、システムは非音紋ベースのデータを

50

監視して適宜識別結果を変えることにより何が間違っているかを判定する。

【0063】

本発明は、追跡目的であらゆる音声コンテンツを識別するために利用できる。デジタル音声ストリーム、アナログ入力またはローカル音声ファイルが全て追跡できる。そのような追跡システムは、音声配信地点に配備されたサーバー側追跡システムであって、報告、デジタル権利管理（DRM）システムまたは権利支払いシステムと統合することができる。追跡されている音声コンテンツが未参加の第三者からのものである場合、本システムのクライアント版を配備して配信中のコンテンツを監視することができる。いずれの場合も、最高水準の精度を保証すべく複数の識別方法を用いる。

【0064】

デジタル権利管理コンポーネントとして波形に基づく識別を利用することが可能であり、ユーザーが作成したデジタル音声ファイルを承認済みコンテンツのリストと比較することが行われる。これにより上述のようなピア・ツー・ピアのファイル共有アーキテクチャ内でフィルター・イン方式が可能になる。

【0065】

音声音紋技術は著作権侵害防止ツールとして用いることができ、調査対象の音声の種類に応じてカスタマイズすることができる。著作権が侵害されたCDの場合、グレースノート社のCDDB CDサービスを利用して目次（TOC）識別を行ない、音声音紋技術を発展させることができる。

【0066】

識別を行うことで、付加価値サービスの配信が可能になる。配信されているコンテンツを明確に把握することができなければ、当該音声コンテンツに関する付加価値コンテンツおよびサービスを配信することは不可能である。

【0067】

本発明の多くの特徴および利点は詳細な明細書から明らかであり、添付の請求項により本発明の真の概念および範囲に含まれるこのような特徴および利点を全て包含することを意図している。更に、当業者は各種の改良や変更を容易に考え付くであろうが、本発明を例示した構成や動作そのものに限定する意図は無く、従ってあらゆる適切な改良および同等物も本発明の範囲に含まれる。例えば、本システムおよび方法は一意な識別子を使用するとして記述しているが、ハッシュ識別子を代わりに用いてもよい。

【図面の簡単な説明】

【0068】

【図1】本発明によるシステムの機能ブロック図である。

【図2】本発明による音紋抽出のフローチャートである。

【図3】未識別録音を識別する方法のフローチャートである。

【図4A】本発明によるシステムのブロック図である。

【図4B】本発明によるシステムのブロック図である。

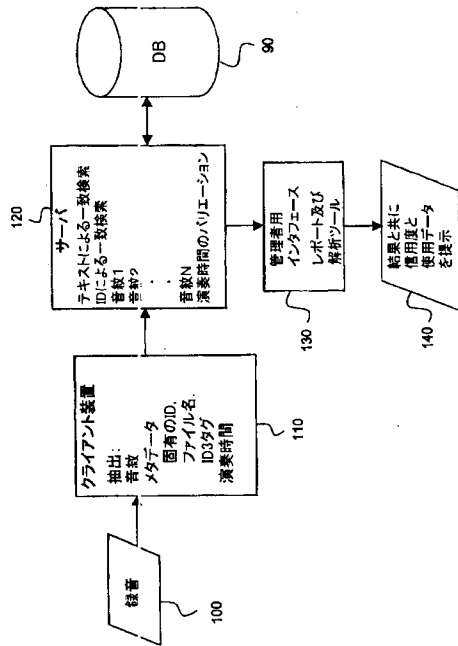
【図4C】本発明によるシステムのブロック図である。

10

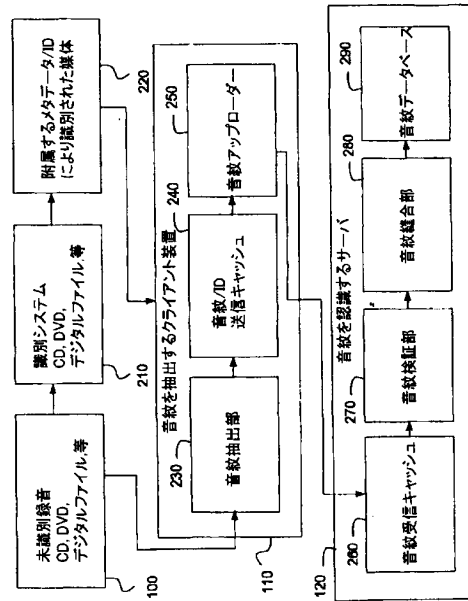
20

30

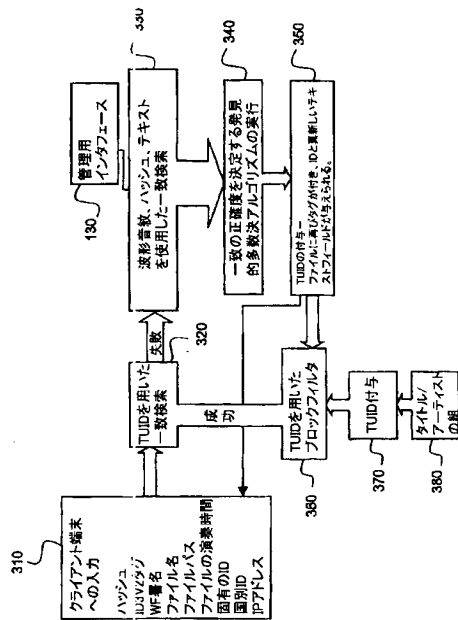
【図 1】



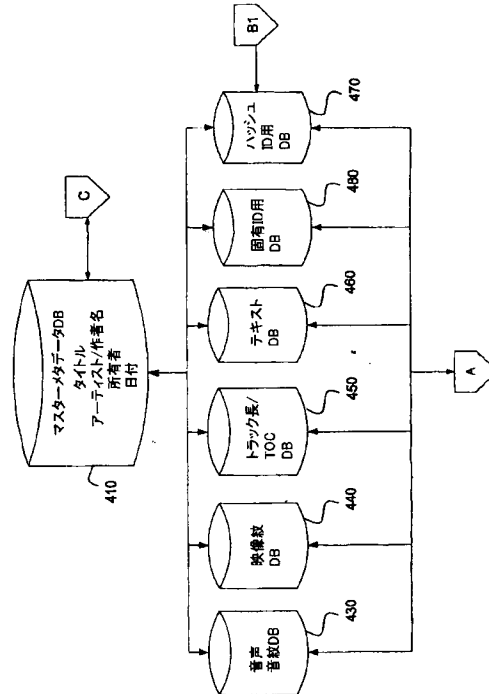
【図 2】



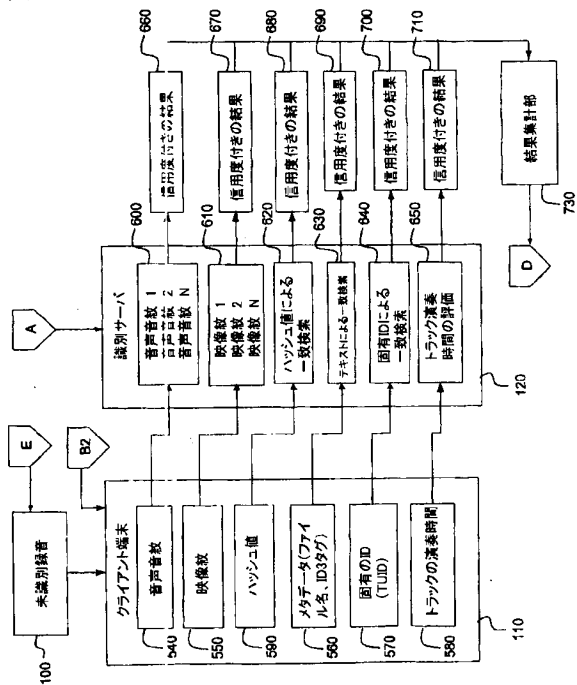
【図 3】



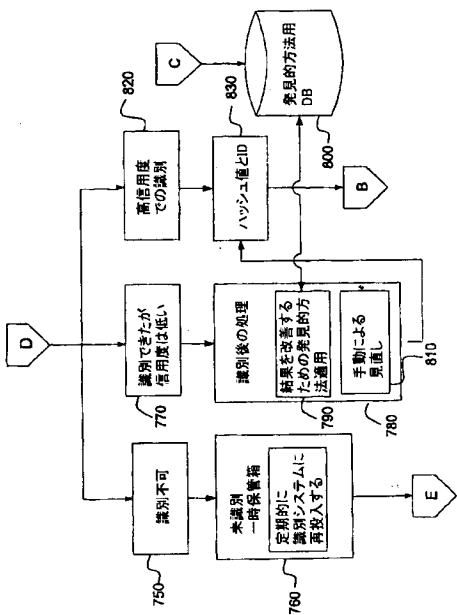
【図 4 A】



【図 4 B】



【図 4 C】



【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(-3) International Publication Date
3 February 2003 (13.02.2003)

PCT

(10) International Publication Number
WO 03/012695 A2

- (51) International Patent Classification: G06F 17/30
- (21) International Application Number: PCT/US02/24054
- (22) International Filing Date: 31 July 2002 (31.07.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/408,594 31 July 2001 (31.07.2001) US
- (71) Applicant: GRACKNOTE, INC. [US/US]; 2141 4th
Street, Berkeley, CA 94710 (US).
- (72) Inventors: ROBERTS, Dale, T.; 15 Oak Springs Drive,
San Anselmo, CA 94960 (US); HYMAN, David; 285
Law Alhaz Drive, Kensington, CA 94708 (US); WHITE,
Stephen; 2526 McKinley Avenue, Berkeley, CA 94703
(US).
- (74) Agent: GOLLHOFER, Richard, A.; Stas & Halsay
L.L., Suite 500, 700 Eleventh Street, N.W., Washington,
DC 20001 (US).
- (81) Designated States (national): AE, AG, AI, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN,
YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GI, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, IE, IT, LI, MC, NL, PT, SE, SK,
TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).
- Published:
— without international search report and to be republished
upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/012695 A2

(54) Title: MULTIPLE STEP IDENTIFICATION OF RECORDINGS

(57) Abstract: Multiple information is extracted from an unknown recording and information associated therewith. Associated information includes the filename, if the recording is a computer file in, e.g., MP3 format, or table of contents (TOC) data, if the recording is on a removable medium, such as a compact disc. At least one and preferably several algorithmically determined fingerprints are extracted from the recording using one or more fingerprint extraction methods. The information extracted is compared with corresponding information in a database maintained for reference recordings. Identification starts with the most accurate and efficient method available, e.g., using a hash ID, a unique ID or text. Fingerprint matching is used to confirm other matches and validation is performed by comparing the duration of the unknown and a possibly matching reference recording.

WO 03/012495

PCT/US02/24054

TITLE OF THE INVENTION

MULTIPLE STEP IDENTIFICATION OF RECORDINGS

CROSS-REFERENCE TO RELATED APPLICATION(S)

[0001] This application is related to and claims priority to U.S. provisional application entitled DIGITAL MUSIC MULTIPLE STEP IDENTIFICATION METHOD AND SYSTEM having serial number 60/308,594, by Dale T. Roberts, et al., filed July 31, 2001, and incorporated by reference herein.

BACKGROUND OF THE INVENTION

1. Field of the Invention

[0002] The present invention is directed to recognition of recordings from their content, and, more particularly to combining fingerprint recognition with other information about a recording to increase reliability of recognition and to accomplish reliable recognition efficiently by using the least expensive forms of recognition first and layering on more complex forms as needed.

2. Description of the Related Art

[0003] There are many uses for recognition of audio (and video) recordings. Many of the uses relate to compensation or control by the rights holders for reproduction and performance of the works recorded. This use of such systems has increased in importance since the development of file sharing software, such as Napster, and the many other similar services available at the end of the twentieth century and the beginning of the twenty first century. Although the need for accurate recognition has been significant for several years, no system has been successful in meeting this need.

[0004] Another use of recording recognition is to provide added value to users when listening (or watching) recordings. One example is the Cddb Music Recognition Service from Gracenote, Inc. of Berkeley, California which recognizes compact discs (CDs) and supplies information regarding a recognized CD, such as album name, artist, track names and access to related content on the Internet, including album covers, artist and fan websites, etc. While the Cddb service is effective for recognizing compact discs, there are several drawbacks in using it to recognize files that are not stored on a removable disc, such as CD or DVD.

[0005] All audio fingerprinting techniques have "blind spots", places where a system using that technique sees similarities and differences in audio where it shouldn't. By relying on just one fingerprinting technique, single source solutions are less accurate when encountering a "blind spot".

WO 03/012695

PCT/US02/24054

[0006] One of the more popular uses for the Gracenote CDDB system is in applications that digitally encode audio files into MP3 and other formats. These encoding applications utilize Gracenote's CDDB service to recognize the compact disc being encoded and to write the correct metadata into the title and ID tags. Gracenote's CDDB service returns a unique ID (TUID) for each track and supports the insertion of such IDs in the ID3V2 tags for MP3 files. The TUID is both hashed and proprietary, and can only be read by the Gracenote system. However, the ID3V2 tags can easily be manipulated to store a TUID for one file in the ID3V2 tag for another file and therefore, the TUID alone is not a reliable identifier of the audio content in a file.

[0007] Gracenote's CDDB service also provides text matching capability that can be utilized to identify digital audio files from their file names, file paths, ID tags (titles), etc. by matching the text extracted by a client device to a metadata database of track, artist, and album names. Although this text matching utilizes user-generated spelling variants associated with each record to improve recognition, there has been no way to verify that the text matches the audio content of the recording once the recording has been separated from a compact disc and stored in a file in any form.

SUMMARY OF THE INVENTION

[0008] An aspect of the present invention is maximizing identification of recordings while minimizing resource usage.

[0009] Another aspect of the present invention is using multiple identification methods so that resource intensive methods, such as audio fingerprinting, are employed only when necessary.

[0010] A further aspect of the invention is minimization of processing of unidentified data.

[0011] Yet another aspect of the present invention is to use the least expensive recognition technique, with progressively more expensive recognition techniques layered onto the process until a desired confidence level is reached.

[0012] A still further aspect of the invention is validation of content-based identification of a recording by comparing text associated with an unidentified recording and text associated with identification records.

[0013] Yet another aspect of the present invention is use of recording identification methods from different sources to increase reliability.

[0014] A still further aspect of the invention is validation of content-based recording identification using fuzzy track length analysis.

W/O 03/012695

PCT/US02/24054

[0015] Yet another aspect of the invention is automatic extraction of identification data for use in a reference database and for identification of recordings.

[0016] A still further aspect of the invention is that unidentified recordings are periodically re-run through the system to determine if recently added data or recently improved techniques will result in recognition.

[0017] The above aspects can be attained by a method of identifying recordings by extracting information about an unknown recording stored in media possessed by a user and at least one algorithmically determined fingerprint from at least one portion of the unknown recording; determining a possible identification of the unknown recording using at least one piece of the information extracted from the unknown recording and an identification database of corresponding information for reference recordings; and identifying the unknown recording when the possible identification based on each of the at least one piece of the information in combination with the at least one algorithmically determined fingerprint identifies a single reference recording with respective confidence levels. The at least one portion of the unknown recording may contain audio, video or both.

[0018] Preferably, the database is maintained by a provider of identification services which supplies unique identifiers that can be recognized only by servers under the control of the provider of identification services. The unique identifiers are associated with recordings once they have been identified. Subsequently, copies of the recordings are recognized using the unique identifiers to greatly speed up the process. The unique identifiers optionally are cached in high-speed RAM or specially indexed database tables.

[0019] When non-waveform data is not available for an unknown recording, the unknown recording is preferably identified by extracting fingerprints from at least one portion of the unknown recording using a plurality of algorithms; determining a possible identification of the unknown recording using at least two of the fingerprints extracted from the unknown recording and at least one database of correspondingly generated fingerprints for reference recordings; and identifying the unknown recording when the possible identification based on each of the fingerprints identifies a single reference recording with respective confidence levels.

[0020] Preferably, an existing database, used to identify recordings possessed by users, which does not contain fingerprint information is expanded by obtaining non-waveform data associated with a recording possessed by a user of the database; extracting at least one fingerprint from at least one portion of the recording; and storing the at least one fingerprint as identifying

WO 03/012695

PCT/US02/24054

information for the recording, when a match is found in the database for the non-waveform data. One example is that during the process of encoding digital music files from an audio CD possessed by a user, a recognition system can be used to identify the audio CD so that fingerprints extracted during the encoding process can be directly associated with the audio CD using a unique ID system.

[0021] Recognition of recordings using either fingerprints or unique identifiers is preferably validated by other information maintained in the identification database, such as the length of the recording or a numeric identifier embedded within the recording. Information about recordings that do not pass validation or match some, but not all of the information used for identification, may be stored for later analysis of the reason for the error. If the fingerprints are obtained as described above, there may have been an error in obtaining the fingerprint. Therefore, errors may be output to an operator, or the system could correct the information stored in the database, based on recognition of patterns in the information that is stored for improper matches. For example, if a large percentage of matching fingerprints are stored, but the other information consistently does not match them, there could be an error in the fingerprint database which needs to be flagged to an operator.

[0022] The present invention includes a system for identifying recordings that includes an extraction unit to extract information about an unknown recording stored in media possessed by a user and at least one algorithmically determined fingerprint from at least one portion of the unknown recording; and an identification unit, coupled to the extraction unit, to make a possible identification of the unknown recording using at least one piece of the information extracted from the unknown recording and an identification database of corresponding information for reference recordings, and to identify the unknown recording when the possible identification based on each of the at least one piece of the information in combination with the at least one algorithmically determined fingerprint identifies a single reference recording with respective confidence levels.

[0023] The present invention also includes a system for identifying recordings that includes an extraction unit to extract fingerprints from at least one portion of an unknown recording using a plurality of algorithms, and an identification unit, coupled to said extraction unit, to make a possible identification of the unknown recording using at least two of the fingerprints extracted from the unknown recording and at least one database of correspondingly generated fingerprints for reference recordings, and to identify the unknown recording when the possible identification

WO 03/012695

PCT/US02/24054

based on each of the fingerprints identifies a single reference recording with respective confidence levels.

[0024] In either of the systems described above, the extraction unit is typically a client unit connected by a network, such as the Internet, to at least one server as the identification unit. The client device may be a personal computer with a drive accessing the recording, a consumer electronics device with a network connection, or a server computer transmitting the unknown recording from one location to another. Furthermore, a portion of the database may be available locally and the extraction unit and identification unit may reside in the same device and share components.

[0025] The present invention also includes a system for obtaining reference information stored in a database used to identify unknown recordings, including a receiving unit to obtain non-waveform data associated with a recording possessed by a user of the database for identification of recordings possessed by the user; an extraction unit to extract at least one fingerprint from at least one portion of the recording; and a storage unit, coupled to said receiving unit and said extraction unit, to store the at least one fingerprint as identifying information for the recording, when a match is found in the database for the non-waveform data.

[0026] These together with other aspects and advantages which will be subsequently apparent, reside in the details of construction and operation as more fully hereinafter described and claimed, reference being had to the accompanying drawings forming a part hereof, wherein like numerals refer to like parts throughout.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a functional block diagram of a system according to the present invention.

Figure 2 is flowchart of a fingerprint extraction according to the present invention.

Figure 3 is a flowchart of a method of recognizing unknown recordings.

Figures 4A-4C are a block diagram of a system according to the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0027] According to the present invention, a suite of identification components are provided in a system like that illustrated in Fig. 1 to facilitate analysis and identification of audio (and video) files utilizing multiple methods. Preferably, an existing database 90 containing recording identifiers and text data is combined with text-based digital audio and audio fingerprinting identification methods. Preferably, the text data in database 90 is obtained from user

WO 03/012695

PCT/US02/24054

submissions and includes user-submitted spelling variants. One such database is available as the CDDB Music Recognition Service from Gracenote, Inc..

[0028] As illustrated in Fig. 1, a recording 100 is accessed by client device 110 via any conventional method, such as reading a digital audio file from a hard drive or a compact disc. Information is extracted from recording 100 and associated information (metadata). Fingerprints are extracted from recording 100, as described in more detail below. The information that is extracted from the metadata includes the duration of the recording which is the track length (from the TOC, for a CD track, the filename and ID3 tag if the recording is in an MP3 file, and the table of contents (TOC) data if the recording is on a CD. If the file containing the recording was produced by a client device operating according to the invention, a unique ID will be extracted from the ID3 file, but initially it will be assumed that information is not available.

[0029] In an exemplary embodiment, the extracted information is sent from client 110 to server 120 to determine a possible identification of the unknown recording using at least one piece of the information extracted from recording 100 and a database 130 of correspondingly generated fingerprints for reference recordings. If text or a unique ID were extracted, an attempt is made to find a match. If a match is found using the text or unique ID, at least one algorithmically determined fingerprint is compared with the fingerprint(s) stored in the matching records to determine whether there is a single reference recording that matches the information extracted from recording 100 with respective confidence levels for each item of information that matches. If no matches can be found based on text and unique ID, an attempt is made to identify the a single reference recording using at least two of the fingerprints extracted from recording 100. If a single reference recording is located using either method, preferably the duration of recording 100 is compared with the duration of the single reference recording as a final validation step.

[0030] Preferably related metadata is used for validation of the match obtained by fingerprint recognition. Like any recognition system fingerprinting can produce erroneous results. Without a validation component such an error can propagate throughout the system and return erroneous data to large percentages of users. The use of validation criteria such as track length comparison enables the system to catch potential errors and flag them for validation.

[0031] A system according to the present invention preferably includes custom result reporting and flexible administrative interfaces 130 to enable weighting of various identification methods and the order of their engagement. Analysis of successful match rates for specific identification methods allows an administrator to manipulate the identifying criteria for each component to

WO 03/012695

PCT/US02/24054

maximize the identification probability. A system according to the present invention preferably incorporates usage data from over 28 million users utilizing the CDDB database via Gracenote Data Services division, to help guide results 140.

[0032] The flexibility of a system according to the present invention allows different configurations to be used for identifying recordings in different environments. An application that monitors streaming audio, for example, requires a very different system and solution architecture than one that identifies files in a peer-to-peer system, or one that identifies analog input. However the present invention can be configured for identification of recordings in each of these situations.

[0033] A system according to the present invention maximizes identification while minimizing resource usage. The use of multiple identification methods ensures that more resource intensive methods, such as audio fingerprinting are employed only when necessary. The use of multiple audio fingerprinting technologies reduces data collision and covers any "blind spots" in a given audio fingerprint technology. The "blind spots" found in single source fingerprinting systems, are avoided by using multiple sources for different fingerprinting techniques. This also provides the ability to fine tune deployment for specific target applications.

[0034] Preferably, fingerprints are obtained using multiple fingerprint recognition services using the method illustrated in Fig. 2. This increases the ability of the system to accurately recognize recordings of various types.

[0035] As illustrated in Fig. 2, when unidentified (unknown) recording 100 is accessed by fingerprint extraction client 110, if possible conventional TOC/file recognition is performed by recognition system 210 and results 220 are returned to fingerprint client 110. Results 220 include a unique identifier (TUID) that points into a master metadata database (not shown in Fig. 2), if the TUID is found. Recording 100 is also processed by fingerprint extractor 230 using at least one and preferably several different algorithmically derived fingerprint extraction systems to obtain fingerprint(s) which are stored in fingerprint/ID send cache 240. As described below in more detail, instructions are received regarding when fingerprint uploader 250 should send the fingerprints to fingerprint recognition server 120.

[0036] In fingerprint recognition server 120, the fingerprints transmitted by fingerprint uploader 250 are initially stored in fingerprint receive cache 260. The fingerprints then undergo fingerprint validation 270 using an algorithmic comparator that attempts to cross-correlate fingerprints for a recording with fingerprints uploaded and extracted by different end users. If it

WO 03/012695

PCT/US02/24054

is found that the fingerprints are substantially similar, they would be validated. This is not the only method that is available for validation, but serves as one example of a process that could be used to reject bad data.

[0037] In this embodiment, fingerprints that are determined to be valid and related undergo stitching 280. For example, if fingerprints are taken from 30 second segments of the recording, the fingerprints are assembled into a continuous fingerprint stream. This could simplify recognition of segments of the recording. The resulting fingerprints are stored in fingerprint database 90 associated with existing database 90 (Fig. 1).

[0038] The CDD3 database has in part been generated through user submissions to create a metadata database with over 12 million tracks and 900,000 albums as of mid-2002. This database contains both basic metadata (artist, album, and track names) as well as extended data (genre, label, etc.).

[0039] A similar distributed collection method may be utilized in the creation of a waveform database using the system illustrated in Fig. 1. In the case where recording 100 is a raw audio waveform, e.g., when a CD is encoded into another format, such as an MP3 file, client device 110 obtains non-waveform data associated with recording 100 which is possessed by a user of database 90 and executes extraction algorithm(s) to extract fingerprints from at least one portion of the recording. The fingerprints are then sent to server 120 with a unique ID, preferably derived from the TOC of the CD. When the unique ID is available, i.e., when a match is found in the database for the non-waveform data, server 120 is able to associate the appropriate metadata in database 90 and the fingerprint(s) with same level of accuracy as identification of CDs by the existing database 90 which is provided for identification of recordings possessed by users. Fingerprints dynamically gathered in this manner may be sent to a fingerprint collection server (not shown in Fig. 1) which would accumulate fingerprints from authenticated clients, as described in more detail below, prior to storing the at least one fingerprint as identifying information for the recording.

[0040] Multiple fingerprint gathering extractors can also be run over a set of static waveforms from a commercial encoder such as Loudeye or Muse. The challenge with this approach is associating the fingerprints with the appropriate metadata. The method described above enables audio fingerprints to be logically associated with parent records and associated back to the original audio source. In the preferred embodiment, the unique ID provides differentiation

WO 03/012695

PCT/US02/24054

between live and studio versions of the same song while simultaneously linking those records to the same artist and their respective albums.

[0041] Preferably server(s) 120 store information in a parallel record set that are linked with unique IDs. When client 110 asks server 120 to recognize media (CD, digital audio file, video file) server 120 may also return a record about how fingerprints should be gathered for this particular CD. This is called the Gathering Instructions Record (GIR). The GIR may include a set of instructions that the remote fingerprint gathering code follows. The record may be pre-computed in off hours or may be dynamically computed at the time of recognition.

[0042] Server 120 may use information it knows about the popularity of a CD to drive decisions about gathering. Everything about a rare CD could be gathered, because the opportunity to get the fingerprints would not want to be missed (even if it was somewhat burdensome to the user). The opposite situation could be true for a very popular CD. The load may be distributed across many users so that they would not even notice that any work for fingerprint gathering was occurring.

[0043] The rules and procedures for building the GIR may be manual, automated and may change over time. They may also be applied uniquely to specific users, applications or geographic locations.

[0044] In one embodiment, the server dynamically gathers fingerprints by modifying the GIR to remove fingerprints that have been gathered previously. The frequency of updating GIRs may vary from instant to delays of days, weeks or months. Some example instructions that may be included in the GIR are:

- A list of track and segments to be gathered and their priority.
- A fingerprint generator algorithm to use.
- Parameters that tell the fingerprint generator how to process the fingerprint, such as:
 - Frequency of audio samples
 - Bands of the frequency domain to process
 - Resolution of the fingerprint
 - Desired Quality of Audio
- When to do the fingerprint gathering, such as
 - Before encoding the track
 - After encoding the track
 - In parallel with encoding the track

WO 03/012695

PCT/US02/24054

- Instructions for caching the fingerprint and when to transmit it back to the server, such as
 - Before encoding the track
 - After encoding the track
 - After the CD has been fully encoded
 - When the communication channel back to the server is not busy
 - When the next CD is loaded up
 - When a group of fingerprints is ready for transmission
- Instructions to take CPU power into the process so as to not overload the computer

[0045] Preferably, the system attempts to improve the quality of the fingerprints during operation. Quality of the source signal, the parameters used for fingerprinting, along with improvements in the fingerprinting algorithms will result in a complex quality matrix that is used by server 120 to determine what fingerprints to gather if higher quality is available. An example of source quality is provided below: Preferably, database 90 or a similar database maintained by fingerprint collection server(s) stores the source quality for fingerprints stored in the database, so that when a fingerprint from higher quality source is available, the fingerprint may be replaced.

Source Quality Table

Name	Bit Rate	Compression	Error Correction	Quality Index
CD Audio Hi-C	44100kbps	None	Hardware	1
CD Audio St-C	44100kbps	None	Software	2
CD Audio	44100kbps	None	None	3
CDR Audio	44100kbps	None	None	4
CDR Made From MP3	44100kbps	mp3	None	5
MP3 File	160kbps	mp3	None	6

[0046] Fingerprints dynamically gathered may contain information that helps validate quality. Information such as errors while reading from the media may be sent up to the fingerprint collector. The system may reject fingerprints that had high error rates from the source media.

[0047] As noted above, instead of immediately storing a fingerprint, multiple fingerprints for a recording may be gathered in by a fingerprint collection server prior to being added to the database. These fingerprints may be compared algorithmically to determine their correlation. If correlation is not adequate then additional fingerprints may be gathered until adequate correlation is achieved and one of the fingerprints or a composite fingerprint is stored in the database. This prevents bad fingerprints from becoming part of the database.

WO 03/012695

PCT/US02/24054

[0048] Stitching of the segmented fingerprints may be necessary since slight variations in timing could result in overlap of the fingerprints. Algorithmic stitching could result in a higher quality continuous fingerprint. Simple stitching appends segmented fingerprints in order of appearance in the recording. Complex stitching could involve scaling different qualities of fingerprints to the lowest common denominator and then appending them in order of their appearance in the recording. Preferably some form of mathematical fitting is utilized if the fingerprint segmentation contains jitter, so that appending is a fuzzy process rather than simple addition of the datastream.

[0049] One example of audio fingerprinting that can be used is described in the U.S. patent application entitled Automatic Identification of Sound Recordings, filed by Maxwell Wells et al. on July 22, 2002 and incorporated herein by reference. However, any known algorithmically derived fingerprinting technique may be used, not only for digital audio, but also video, TV programs (both analog and digital) and DVDs. Appropriate identifiers and recognition techniques will be used for the media to be recognized in a particular application.

[0050] The present invention provides great flexibility and can be utilized for a wide variety of environments, including MP3 recognition in a peer-to-peer environment, or identification of an audio stream for monitoring and reporting purposes. No other solution is known to use multiple recognition components; so it is the only solution that can be customized to meet the needs of any audio (or video) recognition application.

[0051] A functional description for a deployment of the present invention in a peer-to-peer application will be described below with reference to Fig. 3. In this embodiment, audio files are identified before providing public access to them, to determine if the files are allowed in the system, a process known as "filter-in".

[0052] Client device 110 (Fig. 1) extracts information 310 (Fig. 3) from an audio file at the time of upload to server 120 (Fig. 1). The extracted information preferably includes non-waveform data, such as a unique ID, ID3 tag, filename text data, track duration, etc. and fingerprint(s) extracted from the recording and sent to server 120 for recognition.

[0053] The initial match 320 is performed against the unique ID, if present. Use of Gracenote's TUID enables a match to be returned with 99.9% accuracy. This is also the least resource intensive recognition method and can achieve very fast recognition rates. If the unique ID is present the system moves to the validation stage. If no unique ID is present the system attempts identification using the next recognition methods 330.

WO 03/012695

PCT/US02/24054

[0054] In this embodiment, text-based identification is tried next, using a metadata database, such as the Gracenote CDDB service which contains over 900,000 albums and over 12 million songs. Text matching utilizes available text, such as the filename, file path or text within the ID3 tag for MP3 files, to provide a set of data from which to attempt recognition. If an acceptable match is returned, the system moves to the validation stage. If a successful match is not returned, the system attempts identification utilizing the next recognition method.

[0055] The next step is fingerprint identification, in this case using audio fingerprints. The fingerprints from an unknown recording are compared to the fingerprints in database 90 for reference recordings, one fingerprint at a time (or in parallel using different processors for different fingerprints). Each fingerprinting technology returns a match and a level of confidence. If a single reference recording has acceptable confidence levels the system moves to the validation stage. If an unsuccessful match is returned the system can, depending on the target application, ask the user for validation of the most likely result or it can return a "no match found" result.

[0056] Validation is a key component to any successful recognition system. Preferably, key file attributes such as the duration of the recording, are used to validate that a file is what the recognition system says it is by comparing an extracted length of the unknown recording with a stored length of the single reference recording.

[0057] Preferably heuristic and voting algorithms 340 are used to determine if a match is what the system says it is. This self-monitoring reduces the possibility that the system returns inaccurate data that pollutes the system. The heuristics may be manually controlled or algorithmically controlled to produce the best match. These heuristics may also be used to determine which recognition techniques to apply and in what sequence.

[0058] The administrator of each application can determine the level of accuracy needed by each stage (or component) of the system, and therefore has explicit control in optimizing the system. For example, if a 90% aggregate match is required the system administrator can use administrative interfaces 130 to adjust the levels of acceptable return to 90% and a successful result will not be generated unless that threshold is met. The administrator can also set result levels for each component. For example, a 99% text match can be required but only an 85% audio fingerprint match.

WO 03/012695

PCT/US02/24054

[0059] Once a successful identification is returned the file will be retagged 350 with the unique ID allowing for population of the file with the correct ID throughout the system. As a result, future identification of the file will require the least resource intensive recognition method.

[0060] The unique ID (TUID) assigned to the file is then matched 360 against a list 370 of TUIDs populated through the submission of Title/Artist pairs 370 by labels, publishers, and content owners of those files allowed in the system. In one embodiment, if the TUID is present in the database, the file is allowed to be shared, but if the TUID is not present in the database, the file is blocked. In another embodiment, if the TUID is present in the database, the file is blocked. Either of these embodiments could be applied to files recognized as they are accessed by a user, or transmitted from one computer to another.

[0061] As illustrated in Fig. 4A, an embodiment of the present invention uses a plurality of related databases. Master metadata database 410 contains information on title, artist/author name, owner name and date. Related databases include audio fingerprint database 430 and video fingerprint database 440 which form fingerprint database 290 (Fig. 2). Also included are track length/TOC database 450, text database 460, and hash ID database 470 and guaranteed unique ID database 480.

[0062] As illustrated in Fig. 4B, when unidentified (unknown) recording 100 is accessed by client device 110, information is extracted, including fingerprints 540, 550, metadata 560 and unique ID 570, if present. In addition, the duration 580 of the recording is determined and a numerical hash 590 is calculated. The extracted fingerprints are compared with fingerprints 600, 610. Similarly, matching 620, 630, 640 is performed on the numerical hash, text and unique ID. If a reference recording is located, validation is performed by comparing the duration of unidentified recording 100 with the duration of the reference recording. Results 680-710 with a level of confidence for each method of comparison is supplied to result aggregator 730.

[0063] If no reference recording is found 750 matching unidentified recording 100, the extracted information 540-590 and results are stored in unrecognized holding bin 780 for periodic resubmission to recognition server 120 (Figs. 2 & 4B). In this embodiment, if a reference recording is located 770 with a low aggregate confidence level, post recognition processing 780 is performed by applying heuristics 790, or a manual review 810, e.g., by presenting one or more possible matches to the user and receiving the user's selection in response. The results of such user selections may be included in the heuristics stored in heuristics database 820. If post recognition processing 780 results in identification of a single reference recording or result

WO 01/012695

PCT/US02/24054

aggregator 730 outputs recognized results 770 with a high aggregate confidence level, the hash ID is generated 810 and sent to hash database 480 and client device 110, so that the hash and unique ID (TUID) can be stored in the ID3 tag, if a file is being created.

[0064] In one embodiment, the system learns by watching errors in repeated attempts at recognition of similar files to improve its results. It also may receive manual stimulus from users who indicate that there are errors in the results. This allows recognition to be continuously validated over time. For example a file could be recognized by a system according to the invention, then over time the system determines that recognition of that file was flawed, and indicates to an operator that there was something wrong. In another embodiment, the system determines what is wrong by monitoring non-fingerprint based data and changing the recognition results accordingly.

[0065] The present invention can be utilized to identify any audio content for tracking purposes. Digital audio streams, analog inputs or local audio files, can all be tracked. Such a tracking system could be a server side tracking system deployed at the point of audio delivery and integrated with a reporting, digital rights management (DRM) system, or rights payment system. If the audio content being tracked was from a non-participating third party a client version of the system may be deployed to monitor the content being distributed. In either case, multiple identification methods would be utilized to ensure the highest rate of accuracy.

[0066] Utilizing waveform recognition as a digital rights management component is possible, and can be deployed to compare user created digital audio files with lists of approved content. This enables a filter-in approach within a peer-to-peer file sharing architecture such as the one described above.

[0067] Audio fingerprinting technologies can be used as an anti-piracy tool, and can be customized to the type of audio being investigated. In the case of pirated CDs, the Gracenote's Cddb CD service may be utilized to provide table of content (TOC) recognition to augment audio fingerprinting technologies.

[0068] Identification is the enabling component to deliver value-added services. Without explicit knowledge of the content being distributed it is impossible to distribute value-added content and services that relates to that audio content.

[0069] The many features and advantages of the invention are apparent from the detailed specification and, thus, it is intended by the appended claims to cover all such features and

WO 03/012695

PCT/US02/24054

advantages of the invention that fall within the true spirit and scope of the invention. Further, since numerous modifications and changes will readily occur to those skilled in the art, it is not desired to limit the invention to the exact construction and operation illustrated and described, and accordingly all suitable modifications and equivalents may be resorted to, falling within the scope of the invention. For example the system and method have been described as using a unique identifier. However, a hashed identifier could be used instead.

WO 03/012695

PCT/US02/24054

CLAIMS

What is claimed is:

1. A method of identifying recordings, comprising:
 - extracting information about an unknown recording stored in media possessed by a user and at least one algorithmically determined fingerprint from at least one portion of the unknown recording;
 - determining a possible identification of the unknown recording using at least one piece of the information extracted from the unknown recording and an identification database of corresponding information for reference recordings; and
 - identifying the unknown recording when the possible identification based on each of the at least one piece of the information in combination with the at least one algorithmically determined fingerprint identifies a single reference recording with respective confidence levels.
2. A method as recited in claim 1,
 - wherein the identification database is maintained by a provider of identification services, and
 - wherein said determining uses a unique identifier from the provider of identification services when the unique identifier is associated with the unknown recording and otherwise uses text associated with the unknown recording when text is associated with the unknown recording.
3. A method as recited in claim 2, further comprising validating said identifying by comparing an extracted length of the unknown recording with a stored length of the single reference recording.
4. A method as recited in claim 3, wherein the text associated with the unknown recording includes a filename of the recording.
5. A method as recited in claim 3, wherein the text associated with the unknown recording includes an ID3 tag for the recording.
6. A method as recited in claim 1, wherein the at least one algorithmically determined fingerprint is extracted from at least one of audio and video information in the at least one portion of the unknown recording.

WO 03/012695

PCT/US02/24054

7. A method as recited in claim 1,
wherein the at least one algorithmically determined fingerprint includes at least two fingerprints, and

wherein said identifying requires each of the at least two fingerprints to identify the single reference recording with respective confidence levels.

8. A method as recited in claim 1, further comprising validating said identifying by comparing an extracted length of the unknown recording with a stored length of the single reference recording.

9. A method as recited in claim 8, further comprising:
repeating said extracting, determining and identifying for a plurality unknown recordings from a plurality users;
monitoring unsuccessful identifications of the unknown recordings; and
detecting a possible error in the identification database from a pattern of errors.

10. A method as recited in claim 9, wherein said monitoring includes receiving information from the users indicating that said identifying was incorrect.

11. A method as recited in claim 9,
wherein said monitoring includes storing the at least one algorithmically determined fingerprint and an identifier of the single reference recording when said validating is not successful, and
wherein said method further comprises indicating the possible error in the identification database when substantially different fingerprints are stored for a single identifier.

12. A method as recited in claim 9, wherein said method further comprises indicating the possible error when the at least one algorithmically determined fingerprint matches one of the reference recordings, but the unknown recording is associated with ID3 Tag information different from that of the one of the reference recordings.

13. A method as recited in claim 9, further comprising correcting the possible error based on the information extracted from the unknown recordings.

WO 03/012695

PCT/US02/24054

14. A method as recited in claim 8, further comprising indicating a possible error in the identification database when the at least one algorithmically determined fingerprint is substantially similar to one of the reference recordings, but substantially different information is extracted from the unknown recording.

15. A method as recited in claim 1, further comprising delivering related data for the unknown recording from a supplemental database to supplement data embedded within the unknown recording for display and user manipulation.

16. A method of identifying recordings, comprising:
extracting fingerprints from at least one portion of an unknown recording using a plurality of algorithms;
determining a possible identification of the unknown recording using at least two of the fingerprints extracted from the unknown recording and at least one database of correspondingly generated fingerprints for reference recordings; and
identifying the unknown recording when the possible identification based on each of the fingerprints identifies a single reference recording with respective confidence levels.

17. A method as recited in claim 16, wherein each fingerprint is extracted from at least one of audio and video information in the at least one portion of the unknown recording.

18. A method as recited in claim 16, further comprising validating said identifying by comparing a length of the unknown recording with a stored length of the single reference recording.

19. A method as recited in claim 18,
wherein said extracting is performed by client equipment possessed by a user,
wherein said determining, identifying and validating are performed by at least one server under control of a provider of identification services, and
wherein said method further comprises:
transmitting a unique identifier associated with the single reference recording from the at least one server to the client equipment after said validating is successful;
and

WO 03/011695

PCT/US02/24054

associating the unique identifier with the unknown recording in the client equipment.

20. A method as recited in claim 19, further comprising:
comparing the unique identifier with a permission list of stored identifiers in the at least one database;
indicating that the recording may be shared if there is a match for the unique identifier in the permission list.

21. A method as recited in claim 19, further comprising:
comparing the unique identifier with a block list of stored identifiers in the at least one database;
indicating that the recording may not be shared if there is a match for the unique identifier in the block list.

22. A method of obtaining reference information stored in a database used to identify unknown recordings, comprising:
obtaining non-waveform data associated with a recording possessed by a user of the database for identification of recordings possessed by the user;
extracting at least one fingerprint from at least one portion of the recording; and
storing the at least one fingerprint as identifying information for the recording, when a match is found in the database for the non-waveform data.

23. A method as recited in claim 22, wherein the non-waveform data indicates the length of the recording.

24. A method as recited in claim 23, wherein the recording is permanently stored on a removable medium and the non-waveform data is derived from table of contents data for the recording.

25. A method as recited in claim 22, wherein the non-waveform data includes text associated with the recording.

WO 03/012695

PCT/US02/24054

26. A method as recited in claim 25, wherein the non-waveform data includes an ID3 tag.
27. A method as recited in claim 26,
wherein the ID3 tag includes encoded information,
wherein the database is maintained by a provider of identification services and
the encoded information is generated under control of the provider of identification services, and
wherein said method further comprises validating the non-waveform data by
decoding the encoded information prior to said storing of the identifying information in the
database.
28. A method as recited in claim 25, wherein the non-waveform data includes a
watermark.
29. A method as recited in claim 25, wherein the non-waveform data includes media
information regarding source media type.
30. A method as recited in claim 29, wherein the media information identifies the source
media type as CD-R.
31. A method as recited in claim 29, wherein the media information identifies the source
media type as CD-DA.
32. A method as recited in claim 29, wherein the media information identifies the source
media type as a digital file.
33. A method as recited in claim 29, wherein the media information identifies the source
media type as a digital versatile disc.
34. A method as recited in claim 25, wherein the text includes a filename of the
recording.
35. A method as recited in claim 25, wherein the text includes a title of the recording.

WO 03/112695

PCT/US02/24054

36. A method as recited in claim 25, wherein the text includes an artist name of a participant in creation of the recording.

37. A method as recited in claim 25, wherein the text includes an album name associated with the recording.

38. A method as recited in claim 22, wherein said obtaining and extracting are performed by client equipment possessed by a plurality of users for different copies of the recording and different users extract different fingerprints from the recording.

39. A method as recited in claim 38, further comprising:
maintaining the database on at least one server under control of a provider of identification services, and
transmitting from the at least one server to the client equipment, extraction instructions on which of the different fingerprints each of the client equipment extracts.

40. A method as recited in claim 39, further comprising:
transmitting the non-waveform data from the client equipment to the at least one server; and
selecting the extraction instructions by the at least one server for said transmitting to the client equipment based on the non-waveform data.

41. A method as recited in claim 40, further comprising updating the extraction instructions based at least in part on frequency of receipt of the non-waveform data for the recording.

42. A method as recited in claim 40, wherein said selecting of the extraction instructions is based at least in part on type of the client equipment receiving the extraction instructions.

43. A method as recited in claim 40, wherein said selecting of the extraction instructions is based at least in part on geographical location of the client equipment receiving the extraction instructions.

WO 03/012695

PCT/US02/24054

44. A method as recited in claim 40, wherein said selecting of the extraction instructions is based at least in part on software operating on the client equipment receiving the extraction instructions.

45. A method as recited in claim 40, further comprising updating the extraction instructions based at least in part on number of users who have supplied the identifying information.

46. A method as recited in claim 40, wherein said selecting of the extraction instructions is based at least in part on quality of the copies of the recording.

47. A method as recited in claim 40, further comprising transmitting the at least one fingerprint from the client equipment to the at least one server at a time specified by the extraction instructions.

48. A method as recited in claim 47, further comprising storing the at least one fingerprint at the client equipment until a specified number of fingerprints are ready for said transmitting.

49. A method as recited in claim 47, wherein said transmitting of the at least one fingerprint occurs when a communication channel with the at least server is available.

50. A method as recited in claim 47, wherein said transmitting of the at least one fingerprint for a first recording accessed by a piece of client equipment occurs with said transmitting of the non-waveform data for a second recording accessed by the piece of client equipment.

51. A method as recited in claim 47,
wherein the recording is permanently stored on a removable medium and the client equipment generates at least one encoded file from the recording, and
wherein said transmitting transmits the at least one fingerprint before encoding the recording.

52. A method as recited in claim 47,

WO 01/012695

PCT/US02/24054

wherein the recording is permanently stored on a removable medium and the client equipment generates at least one encoded file from the recording, and
wherein said transmitting transmits the at least one fingerprint after encoding one track of the removable medium.

53. A method as recited in claim 47,
wherein the recording is permanently stored on a removable medium and the client equipment generates at least one encoded file from the recording, and
wherein said transmitting transmits the at least one fingerprint after receiving an indication that encoding of the removable medium has been completed.

54. A method as recited in claim 22, wherein the database includes the identifying information for musical recordings.

55. A method as recited in claim 22, wherein the database includes the identifying information for video recordings.

56. A method as recited in claim 22, further comprising:
detecting a quality of the at least one fingerprint;
identifying another copy of the recording using the at least one fingerprint; and
replacing the at least one fingerprint with a higher quality fingerprint when the other copy of the recording produces the higher quality fingerprint.

57. A method as recited in claim 56, wherein said detecting of the quality is based on an encoding technique used for the recording.

58. A method as recited in claim 56, wherein said detecting of the quality is based on a media type used to store the recording.

59. A method as recited in claim 56, wherein said detecting of the quality is based on error correction capability of user equipment accessing the recording.

WO 03/012695

PCT/US02/24054

60. A method as recited in claim 59, wherein said detecting of the quality assigns higher quality when hardware error correction is used than when software error correction is by the user equipment.

61. A method as recited in claim 56, wherein said detecting of the quality is based on number of errors detected during said extracting of the fingerprint.

62. A method as recited in claim 22,
 wherein said obtaining and extracting are performed by client equipment possessed by a plurality of users for different copies of the recording, and
 wherein said method further comprises:
 comparing the at least one fingerprint obtained from one of the users with the at least one fingerprint extracted from at least one other user; and
 updating the at least one fingerprint in the database based on said comparing.

63. A method as recited in claim 62, wherein said updating is performed after said comparing determines that fingerprints from different users have a predetermined correlation.

64. A method as recited in claim 62, wherein said updating combines fingerprints from different users for storage in the database.

65. A system for identifying recordings, comprising:
 an extraction unit to extract information about an unknown recording stored in media possessed by a user and at least one algorithmically determined fingerprint from at least one portion of the unknown recording; and
 an identification unit, coupled to said extraction unit, to make a possible identification of the unknown recording using at least one piece of the information extracted from the unknown recording and an identification database of corresponding information for reference recordings, and to identify the unknown recording when the possible identification based on each of the at least one piece of the information in combination with the at least one algorithmically determined fingerprint identifies a single reference recording with respective confidence level.

WO 03/012695

PCT/US02/24054

66. A system for identifying recordings, comprising:
an extraction unit to extract fingerprints from at least one portion of an unknown recording using a plurality of algorithms; and
an identification unit, coupled to said extraction unit, to make a possible identification of the unknown recording using at least two of the fingerprints extracted from the unknown recording and at least one database of correspondingly generated fingerprints for reference recordings, and to identify the unknown recording when the possible identification based on each of the fingerprints identifies a single reference recording with respective confidence levels.

67. A system for obtaining reference information stored in a database used to identify unknown recordings, comprising:
a receiving unit to obtain non-waveform data associated with a recording possessed by a user of the database for identification of recordings possessed by the user;
an extraction unit to extract at least one fingerprint from at least one portion of the recording; and
a storage unit, coupled to said receiving unit and said extraction unit, to store the at least one fingerprint as identifying information for the recording, when a match is found in the database for the non-waveform data.

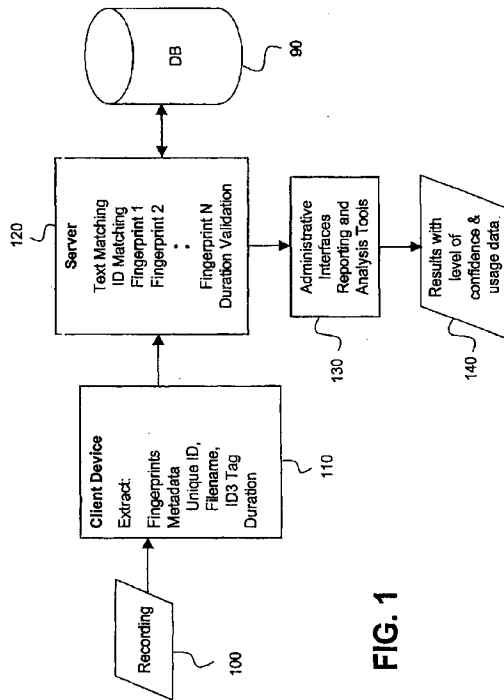


FIG. 1

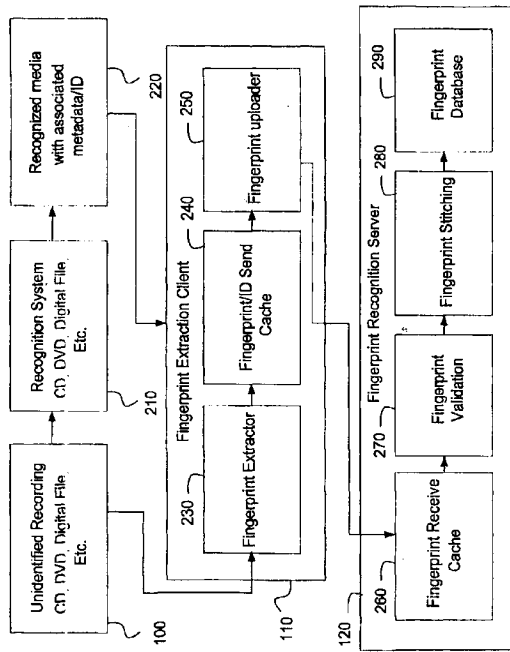
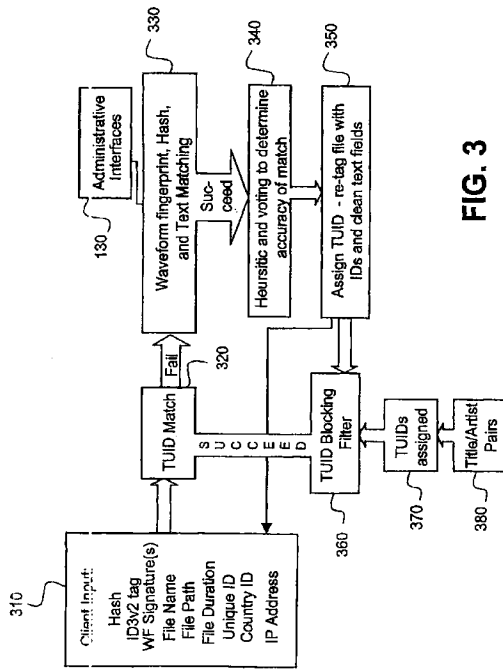
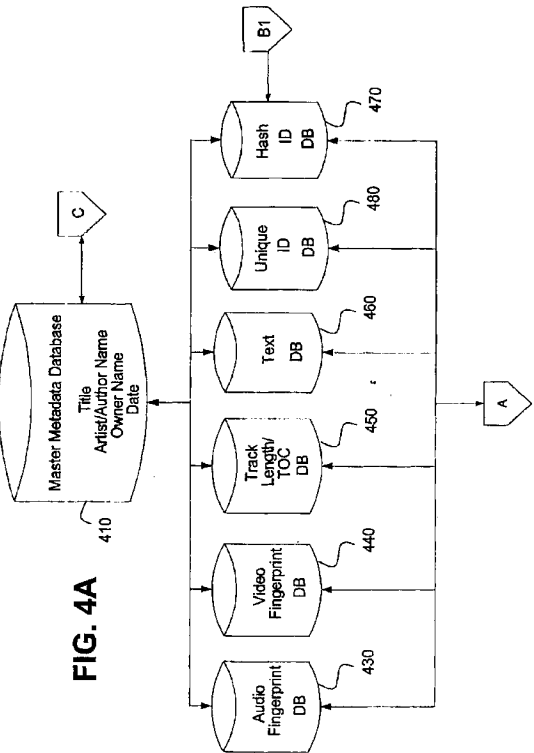


FIG. 2





WO 03/012695

5/6

PCT/US02/24054

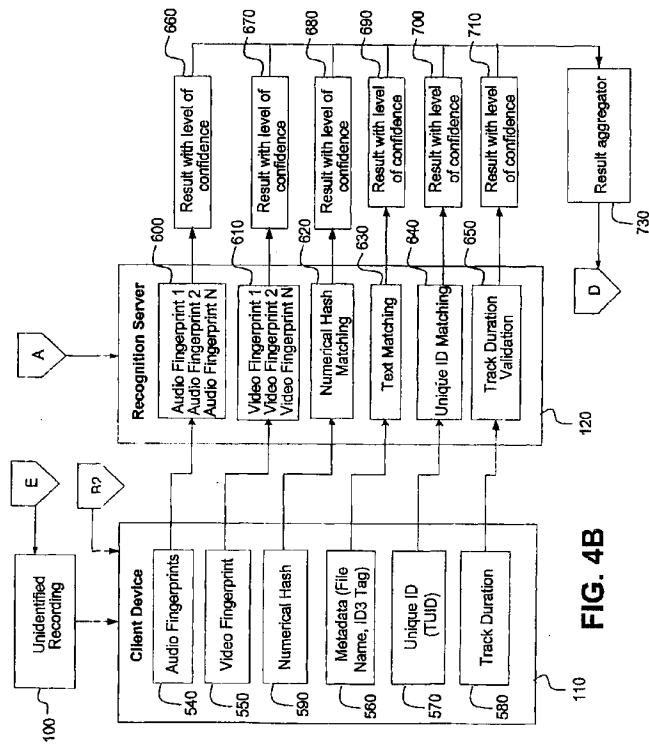


FIG. 4B

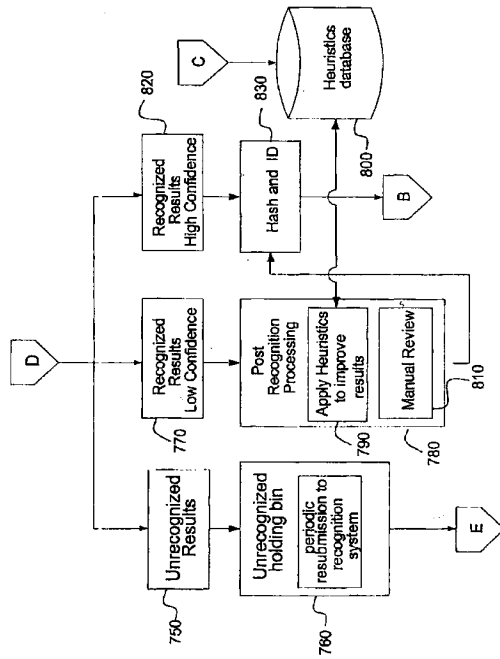


FIG. 4C

【国際公開パンフレット（コレクトバージョン）】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

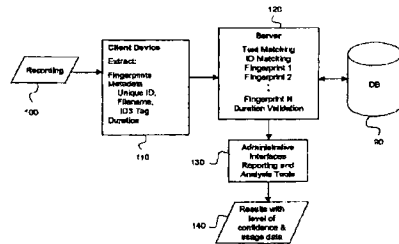
(19) World Intellectual Property
Organization
International Bureau(43) International Publication Date
13 February 2003 (13.02.2003)

PCT

(10) International Publication Number
WO 2003/012695 A3

- (51) International Patent Classification: G06F 17/30 (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GR, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.
- (21) International Application Number: PCT/US2002/024054 (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW); Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM); European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR); OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (22) International Filing Date: 31 July 2002 (31.07.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/408,594 31 July 2001 (31.07.2001) US
- (71) Applicant: GRACENOTE, INC. [US/US]; 2141 4th Street, Berkeley, CA 94710 (US).
- (72) Inventors: ROBERTS, Dale, T.; 15 Oak Springs Drive, San Anselmo, CA 94960 (US); HYMAN, David; 285 Los Altos Drive, Kensington, CA 94708 (US); WHITE, Steven; 2326 McKinley Avenue, Berkeley, CA 94703 (US).
- Published: with international search report
- (86) Date of publication of the international search report: 18 March 2004
- (74) Agent: GOLLHOFER, Richard, A.; Stas & Halsey LLP, Suite 500, 700 Eleventh Street, N.W., Washington, DC 20001 (US).
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) TITLE: MULTIPLE STEP IDENTIFICATION OF RECORDINGS



(57) Abstract: Multiple information is extracted from an unknown recording and information associated therewith. Associated information includes the filename, if the recording is a computer file in, e.g., MP3 format, or table of contents (TOC) data, if the recording is on a removable medium, such as a compact disc. At least one and preferably several algorithmically determined fingerprints are extracted from the recording using one or more fingerprint extraction methods. The information extracted is compared with corresponding information in a database maintained for reference recordings. Identification starts with the most accurate and efficient method available, e.g., using a hash ID, a unique ID or text. Fingerprint matching is used to confirm other matches and validation is performed by comparing the duration of the unknown and a possibly matching reference recording.

WO 2003/012695 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 02/24054
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 H06F17/30		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 H06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 319 567 B (MCS PARTNERS) 14 June 1989 (1989-06-14)	1, 6, 7, 15-17, 65, 66 2, 15
Y	page 2, line 3, 5, 28, 29 page 3, line 22 page 4, line 17 - page 5, line 4 page 5, line 46 - page 6, line 32 page 8, line 30 --- -/-	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "A" document member of the same patent family		
Date of the actual completion of the international search 19 November 2003		Date of mailing of the international search report 15/12/2003
Name and mailing address of the ISA European Patent Office, P.O. 5010 Patenkamp 2 14, - 2280 HV Rijswijk Tel (+31-70) 340-2048, Tx 31 661 epo nl, Fax (+31-70) 340-3016		Authorized officer Bykowski, A

Form PCT/ISA/210 (in word sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 02/24054

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 01 28222 A (TANG JIMMY TAI KWAN ; LAI SAI KIT (CN); NG KIN PING (CN); TSUI CHI) 19 April 2001 (2001-04-19)	22-26, 28-37, 54,55 2,15
Y	page 1, line 8-10 page 4, line 8-24 page 5, line 20 -page 6, line 2 page 15, line 9 -page 16, line 5 page 17, line 6-8 page 28, line 21,22 page 30, line 1 -page 32, line 29	
X	WO 98 25269 A (THOMSON CONSUMER ELECTRONICS) 11 June 1998 (1998-06-11) page 3, line 12 -page 4, line 37 page 5, line 22 -page 6, line 11 figures 3,4 page 8, line 8-12	1,6,7, 15-17, 65,66
X	"November 2000" THE MUSICBRAINZ MAILING LIST ARCHIVE, 'Online! December 2000 (2000-12), XP002262122 Retrieved from the Internet: <URL:http://www.musicbrainz.org/pipermail/ musicbrainz/2000-November.txt.gz> 'retrieved on 2003-11-15!	22-64,67
A	page 16, line 14,15 page 38, line 17,18 page 65, line 3-8 page 77, line 24-32 page 83, line 15 -page 86, line 10 page 88, line 7-25	1-22,65, 66
A	US E 918 223 A (KEISLAR DOUGLAS F ET AL) 29 June 1999 (1999-06-29) column 1, line 61-66 -column 2, line 51-60 column 3, line 56-67 column 5, line 53 -column 6, line 10 column 6, line 43,44 column 15, line 30-34 column 17, line 20,21 column 22, line 10-16	1-67

Form PCT/ISA210 (continuation of second sheet) (July 2002)

INTERNATIONAL SEARCH REPORT
 Information on patent family members

 International Application No.
PCT/US 02/24054

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0319567	B	14-06-1989	US 4843562 A
			AT 85735 T
			AU 2075288 A
			AU 613941 B2
			BR 8807105 A
			CA 1290063 C
			DE 3878423 D1
			DE 3878423 T2
			EP 0319567 A1
			ES 2012834 A6
			HK 1000982 A1
			IL 86701 A
			JP 2500875 T
			JP 3001896 B2
			KR 9600152 B1
			MX 166116 B
			WO 8810540 A1
WO 0128222	A	19-04-2001	AU 2018201 A
			WO 0128222 A2
WO 9825269	A	11-06-1998	US 6034925 A
			US 6031795 A
			AU 7627398 A
			CN 1284721 A
			CN 1284722 A
			CN 1239573 A
			DE 69716632 D1
			DE 69716632 T2
			EP 1152421 A2
			EP 0941537 A1
			JP 2001506795 T
			JP 2001313892 A
			KR 2000057236 A
			WO 9825269 A1
			US 6252831 B1
			US 6226235 B1
			JP 2001273748 A
US 5318223	A	29-06-1999	NONE

Form PCT/ISA210 (as amended) (July 1999)

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,SK,TR),OA(BF,BJ,CF,CG,CI,CM,CA,CN,GQ,GW, ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES, FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,N O,NZ,OM,PH,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZM,ZW

(72)発明者 デイビット ハイマン

アメリカ合衆国 カリフォルニア州 94708 ケンジントン ロスアルトスドライブ 285

(72)発明者 ステファン ホワイト

アメリカ合衆国 カリフォルニア州 94703 パークレイ マクキンリーアヴェニュー 23
26

Fターム(参考) 5D015 AA03 BB02 CC17 KK02 LL07

(54)【発明の名称】録音の多段階識別関連出願の相互参照本出願は、「デジタル音楽の多段階識別の方法およびシステム」と題された米国仮出願第60/308,594号(発明者:デールT. ロバーツ(Dale T. Roberts))ら、出願日:2001年7月31日)に関連し、これを優先権主張するとともに、本願明細書に援用するものである。